Investigation in South Africa of viruses implicated in a maize lethal necrosis disease outbreak in East Africa

by

Natalie Nel

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in the Faculty of Science at Stellenbosch University.

Supervisor: Prof. Gerhard Pietersen

March 2021

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third-party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Natalie Nel
March 2021

Abstract

Virus diseases of maize (*Zea mays*) such as maize streak virus (MSV) and the recently identified maize lethal necrosis disease (MLND) may result in severe to complete maize yield losses for individual farmers in sub-Saharan Africa in any given year, threatening food security. MSV has been reported as widespread in South Africa since the 1870s, while MLND is yet to be reported in the country due to the current absence of one of the primary viruses required for MLND expression, namely maize chlorotic mottle virus (MCMV). Maize in South Africa may be pre-disposed to MLND as maize-infecting potyviruses, required for synergistic coinfection with MCMV to cause MLND expression, have been reported in South African maize previously, along with the major known vectors of MCMV and potyviruses: thrips and aphids. Furthermore, South Africa's climate is ideal for both MCMV and its other known vectors to thrive should they be introduced, with KwaZulu-Natal being one of the provinces most at risk. MCMV is predicted to spread into South Africa through Mozambique and/or Zimbabwe. To better understand the risk of a MLND outbreak occurring in South Africa, maize grown in KwaZulu-Natal was surveyed using polymerase chain reaction (PCR) for viruses recently implicated in a MLND outbreak in Tanzania. These viruses included MCMV, potyvirids, MSV, maize-associated pteridovirus (MaPV), Morogoro maize-associated virus (MMaV) and two maize-associated totivirus (MATV) variants. Representatives of samples containing viruses not reported in South Africa previously were analysed with next generation sequencing (NGS). Furthermore, the genetic diversity of MSV was also determined across other major maize-growing regions in South Africa as the current state of this important virus was unknown, with the most recent previous study on MSV in the country conducted on plants sampled over 20 years ago. No infections of MCMV or potyvirids were detected in maize during this study. However, the presence of MaPV and MMaV was detected and confirmed for the first time in South Africa, as well as maize stripe virus (MStV) whose presence in the country, although reported, is not based on any published account. Other viruses, such as two MATV variants, maize streak Reunion virus (MSRV), and two strains of Zea mays chrysovirus 1, are regarded as preliminary findings in this study as their detection was not pursued further. MSV continues to be widespread in the country, with hotspots detected in the Pongola region of KwaZulu-Natal, Eswatini and the Ofcolaco region of Limpopo. The current genetic diversity of MSV present in South Africa appears similar to that described 20 years ago.

Opsomming

Verskeie virussiektes van mielies (*Zea mays*), insluitend mieliestreepvirus (MSV) en die onlangs geïdentifiseerde siekte, mieliedodelike nekrose (MLND), kan lei tot ernstige tot algehele opbrengs verliese vir individuele boere in Afrika suid van die Sahara in enige gegewe jaar, en bedreig dus voedsel sekuriteit. MSV is al sedert die 1870's wydverspreid in Suid Afrika gerapporteer, terwyl MLND nog nie in die land gerapporteer is nie weens die huidige afwesigheid van een van die primêre virusse wat benodig word vir MLND-uitdrukking, naamlik mielieschlorotiese vlekvirus (MCMV). Mielies in Suid Afrika is egter vatbaar vir MLND aangesien mieliebesmettende potyvirusse, die ander komponent wat benodig word vir MLND-uitdrukking, reeds in Suid Afrikaanse mielies gerapporteer is. Die belangrikste vektore van MCMV en potyviruses, naamlik blaaspootjies en verskei mielie plantluise ko mook reeds hier voor. Verder is die klimaat van Suid Afrika ideaal vir beide MCMV en van die ander gerapporteerde vektore om te floreer sou MCMV die land binnekom. Daar word voorspel dat MCMV bes moontlik deur Mosambiek en/of Zimbabwe na Suid Afrika sal versprei, met KwaZulu Natal as een van die provinsies is wat die grootste in gevaar is. Om die risiko van 'n MLND-uitbraak in Suid Afrika te ondersoek, is mielies wat in KwaZulu Natal verbou is, met behulp van polimerase kettingreaksie (PCR) getoets vir die virusse wat onlangs tydens 'n MLND-uitbraak in Tanzanië betrokke was. Hierdie virusse sluit in MCMV, potyvirids, MSV, mielie-geassosieerde pteridovirus (MaPV), Morogoro mielie-geassosieerde virus (MMaV) en twee mielie-geassosieerde totivirus (MATV) variante. Verteenwoordigende monsters is met die volgende generasie volgordebepalings (NGS) geanaliseer om te bepaal of hulle virusse bevat wat nie voorheen in Suid Afrika gerapporteer is nie,. Verder is die genetiese diversiteit van MSV ook bepaal dwarsdeur belangrike mielie-groeiende streke in Suid Afrika bepaal. Dit is gedoen aangesien die huidige toestand van hierdie belangrike virus onbekend was, met die mees onlangse vorige studie wat op monsters gedoen is wat 20 jaar gelede versamel was. Geen infeksies van MCMV of potiviriede is in mielies tydens die huidige studie opgespoor nie. MaPV en MMaV is egter vir die eerste keer in Suid Afrika gevind, asook mieliestreepvirus (MStV) wat wel vroeer in die gerapporteer was, maar wat nie gebaseer was op n gepubliseerde rekord nie. Ander virusse, soos twee MATV-variante, mieliestreep Reunion-virus (MSRV), en twee stamme van Zea mays chrysovirus 1, word in hierdie studie as voorlopige bevindings beskou, aangesien die opsporing daarvan nie verder nagestreef is nie. MSV is steeds wydverspreid in die land, met brandpunte in die Pongola-streek in KwaZulu-Natal, Eswatini en die Ofcolaco-streek in Limpopo. Die huidige genetiese diversiteit van MSV wat in Suid Afrika voorkom, lyk soortgelyk aan die wat 20 jaar gelede beskryf is.

Acknowledgements

I would like to acknowledge the following people and institutions, whom without this research would not have been possible:

- My supervisor, Prof. Gerhard Pietersen, for his guidance, giving me the opportunity to do this research, enabling and encouraging me to attend and present at my first international conference, and for always believing in and supporting me.
- Dr Barbara van Asch for her guidance relating to the various phylogenetic analysis tools employed during this study.
- Dr David A. Read for supplying and performing the numerous NGS runs, free of charge.
- Gert Pietersen for assisting with sampling and PCR testing for maize chlorotic mottle virus and potyvirids.
- Abraham Twala for assisting with sample collection.
- Dr Beatrix Coetzee for her mentorship, friendship, and endless support.
- My friends and colleagues in the Vitis lab that provided input into this project and kept morale high.
- The South African National Seed Organization (SANSOR) for both personal and project-based financial assistance. Opinions expressed and conclusions arrived at are those of the author and not necessarily to be attributed to SANSOR.
- Stellenbosch University.
- Kyle J. Koekemoer for being my home away from home, and for his love, patience, positivity, and emotional support throughout this study.
- Brandon Nel for being there for me in times of need, and times of tea.
- My grandparents for their love and encouragement.
- My parents for always believing in me and for their endless love, encouragement, and financial support throughout my academic career.
- My Heavenly Father.

Table of contents

List of abbreviations

%	Percentage
°C	Degrees Celsius
+F	Frequencies
+G	Gamma distribution
+I	Evolutionary invariant sites
µM	Micromolar
µl	Microlitre
6K	Six kilodalton peptide
aa	Amino acid
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
BLASTn	BLAST (search a nucleotide database using a nucleotide query)
bp	Base pairs
cDNA	Complementary deoxyribonucleic acid
CDS	Coding domain sequence
CH	Switzerland
CI	Cytoplasmic inclusion
CIMMYT	International Maize and Wheat Improvement Center
CLND	Corn lethal necrosis disease
cm	Centimetre
CP	Coat protein
CTAB	Cetyltrimethylammonium bromide
dNTP	Deoxyribonucleotide triphosphate
DNA	Deoxyribonucleic acid
DRC	Democratic Republic of the Congo
DSV	Digitaria streak virus
DTT	Dithiothreitol
E-value	Expectation value
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EPPO	European and Mediterranean Plant Protection Organization
FAO	Food and Agriculture Organization
FAO REOA	Food and Agriculture Organization Regional Emergency Coordinator for Eastern Africa
FAOSTAT	Statistical databases and datasets of the FAO
g	Grams
GPS	Global Positioning System

HC-Pro	Helper component-proteinase
ICTV	International Committee on Taxonomy of Viruses
JGMV	Johnsongrass mosaic virus
KARLO	Kenya Agricultural and Livestock Research Organization
kb	Kilobase pairs
kcal	Kilocalorie
LIR	Long intergenic region
M	Molar
mg	Milligrams
mM	Millimolar
MaPV	Maize-associated pteridovirus
MaYDV-RMV	Maize yellow dwarf virus-RMV
MaYMV	Maize yellow mosaic virus
MATV	Maize-associated totivirus
MCMV	Maize chlorotic mottle virus
MDMV	Maize dwarf mosaic virus
MEGA X	Molecular Evolutionary Genetic Analysis X
min	Minute
MLND	Maize lethal necrosis disease
MMaV	Morogoro maize-associated virus
MSc.	Master of Science
MSRV	Maize streak Reunion virus
MSV	Maize streak virus
MStV	Maize stripe virus
MP	Movement protein
mRNA	Messenger ribonucleic acid
MUSCLE	Multiple Sequence Comparison by Log-Expectation
NaCl	Sodium chloride
NCBI	National Centre of Biotechnology Information
NGS	Next generation sequencing
NIa-Pro	Nuclear inclusion A protease
NIb	Nuclear inclusion B body
Nm	Nanometres
nt	Nucleotides
OTU	Operational taxonomic unit
P1-Pro	Protein 1 protease
PCR	Polymerase chain reaction
PIPO	Pretty interesting *Potyviridae* open reading frame
PVP	Polyvinyl pyrrolidone
RdRp	RNA-dependant RNA polymerase

Rep	Replicase-associated protein
RepA	Replicase-associated protein A
RNA	Ribonucleic acid
RSV	Rice stripe virus
RT-PCR	Reverse transcription-polymerase chain reaction
s	Seconds
SANSOR	South African National Seed Organization
SCMV	Sugarcane mosaic virus
SIR	Short intergenic region
TBE	Tris-borate-EDTA
TPCTV	Tomato pseudo-curly top virus
Tris-HCl	Tris (hydroxymethyl) aminomethane hydrochloride
U	Unified atomic mass units
USA	United States of America
V	Volts
WSMV	Wheat streak mosaic virus
X	Fold
ZA	South Africa
ZMCV1	Zea mays chrysovirus 1

List of figures

List of tables

Introduction

Background

Maize (*Zea mays*) provides a staple food source and form of livelihood to more than 300 million people in sub-Saharan Africa (Edmeades 2008). Unlike the rest of the world, in Africa, white maize is grown preferentially to yellow maize, contributing around 90% of the total maize produced in Africa each year, and around 30% of the yellow maize produced globally (Ekpa et al. 2018; Khumalo et al. 2011). Viral pathogens threaten the maize industry with severe to complete losses attributed to maize streak disease (MSD) (Bosque-Pérez 2000). Major losses have also been reported for plants expressing maize lethal necrosis disease (MLND) (Mahuku et al. 2015; Pratt et al. 2017; Redinbaugh and Stewart 2018), which recently emerged in sub-Saharan Africa (Wangai et al. 2012) where it spread rapidly (Redinbaugh and Stewart 2018).

Maize streak virus (MSV), the causative agent of MSD, has been reported in maize in South Africa for over 100 years (Fuller 1901). MLND is caused by the co-infection of maize chlorotic mottle virus (MCMV) and any maize-infecting potyvirid, with MCMV considered as the virus responsible for MLND emergence (Redinbaugh and Stewart 2018). In 2019, a variety of other viruses, including MSV, were detected in MLND-affected plants (Read et al. 2019a, b, c, d), and are expected to potentially contribute to disease severity (Redinbaugh and Stewart 2018), although this is yet to be confirmed. Southward expansion of MCMV in sub-Saharan Africa has been observed, suggesting the virus may spread to South Africa (Isabirye and Rwomushana 2016).

The South African maize industry is concerned that the introduction of this disease into the country may be devastating for both commercial and smallholder farmers and for the country's economy. Thus, developing effective methods of disease prevention, conducting routine virus surveillance, and determining the risk of a MLND outbreak should MCMV be introduced into the country, are currently areas of high priority.

Problem statement

The current state of viruses affecting maize in South Africa is unknown, therefore making it difficult to accurately determine the country's possible predisposition to MLND and the risk of a possible outbreak occurring should an incursion of MCMV occur.

Aims and objectives

The aim of this study was to determine the incidence and distribution of MLND-implicated viruses in KwaZulu-Natal, South Africa, and to determine the current genetic diversity and distribution of MSV across the major maize-growing regions of the country. To achieve this aim, the following objectives were formulated:

- To sample maize plants with virus-like symptoms along KwaZulu-Natal's major maize grain transport route, and samples with MSV-like symptoms along four other major maize grain transport routes in South Africa.

- To obtain total nucleic acid from the samples of high concentration and purity.
- To use polymerase chain reaction (PCR) to detect infections of MLND-implicated viruses including MCMV, potyvirids, MSV, maize-associated pteridovirus, Morogoro maize-associated virus and two maize associated totivirus variants.
- To confirm the presence of viruses new to South Africa using a second detection system, namely next generation sequencing (NGS).
- To amplify a hypervariable region of the MSV genome using PCR.
- To use phylogenetic tools to assess the genetic diversity of MSV present in the samples and the geographical distribution thereof.
- To assemble whole genome sequences for MSV variants present in representative samples selected from the phylogenetic analyses of the PCR results.
- To determine the whole genome-based genetic diversity of MSV present within the selected samples.
- To determine the presence of any novel viruses detected in the samples selected for NGS.

Research outputs

MSV findings were presented in the form of an oral presentation at an international conference, Virology Africa, in Cape Town on 11 February 2020. The findings of the survey of maize in KwaZulu-Natal along with the first report of MaPV and MMaV have been submitted to the European Journal of Plant Pathology for possible publication. The detection of MSRV and the detection and confirmation of MStV have prompted further research, which is currently being conducted, to confirm the genome sequences of these virus variants, after which the presence of these viruses in South Africa will be reported in the form of an article in a peer-reviewed journal, possibly accompanied by genome announcements should the variants be different enough from those previously reported. A total of thirteen complete/near complete virus sequences were uploaded to GenBank with the following accessions: MATV: MW063115 and MW063116; MaPV RNA1 and RNA2: MW063117 and MW063118, respectively; MStV RNA1-5: MW063119-MW063123, respectively; MMaV: MW063124; ZMCV1-63: MW063135; and ZMCV1-201: MW063136.

References

Bosque-Pérez, N. A. (2000). Eight decades of maize streak virus research. *Virus Research*, 71(1–2), 107–121.

Edmeades, G. O. (2008). Drought tolerance in maize: An emerging reality. Resource document. International Service for the Acquisition of Agri-Biotech Applications (ISAAA).

Ekpa, O., Palacios-Rojas, N., Kruseman, G., Fogliano, V., & Linnemann, A. R. (2018). Sub-Saharan African maize-based foods: Technological perspectives to increase the food and nutrition security impacts of maize breeding programmes. *Global Food Security*, 17, 48–56.

Fuller, C. (1901). Mealic variegation. *First Report of the Government Entomologist Natal 1899-1900* (pp. 17–19). Pietermaritzburg: P. Davis & Sons, Government Printers.

Isabirye, B. E., & Rwomushana, I. (2016). Current and future potential distribution of maize chlorotic mottle virus and risk of maize lethal necrosis disease in Africa. *Journal of Crop Protection*, 5(2), 215–228.

Khumalo, T. P., Schönfeldt, H. C., & Vermeulen, H. (2011). Consumer acceptability and perceptions of maize meal in Giyani, South Africa. *Development Southern Africa*, 28(2), 271–281.

Mahuku, G., Lockhart, B. E., Wanjala, B., Jones, M. W., Kimunye, J. N., Stewart, L. R., et al. (2015). Maize lethal necrosis (MLN), an emerging threat to maize-based food security in sub-Saharan Africa. *Phytopathology*. https://doi.org/10.1094/PHYTO-12-14-0367-FI

Pratt, C. F., Constantine, K. L., & Murphy, S. T. (2017). Economic impacts of invasive alien species on African smallholder livelihoods. *Global Food Security*, 14(November 2016), 31–37.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019a). Diversity and distribution of maize-associated totivirus strains from Tanzania. *Virus G*enes, 55, 429–432.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019b). Molecular characterization of Morogoro maize-associated virus, a nucleorhabdovirus detected in maize (Zea mays) in Tanzania. *Archives of Virology*, 164, 1711–1715.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019c). Characterization and detection of maize-associated pteridovirus (MaPV), infecting maize (Zea mays) in the Arusha region of Tanzania. *European Journal of Plant Pathology*, 154, 1165–1170.

Read, D. A., Featherstone, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019d). First report of maize yellow mosaic virus (MaYMV) on maize (Zea mays) in Tanzania. *Journal of Plant Pathology*, 101(1), 203.

Redinbaugh, M. G., & Stewart, L. R. (2018). Maize lethal necrosis: An emerging, synergistic viral disease. *Annual Review of Virology*, 5(August), 301–322.

Wangai, A. W., Redinbaugh, M. G., Kinyua, Z. M., Miano, D. W., Leley, P. K., Kasina, M., et al. (2012). First report of maize chlorotic mottle virus and maize lethal necrosis in Kenya. *Plant Disease*, 96(10), 1582.

Chapter 1: Literature review

1.1 The maize industry

It is believed that maize (*Zea mays* L.), also known as corn, was domesticated around 3,000-8,000 years ago (Smith 1995; Wang et al. 1999) from a wild Mexican grass teosinte (*Zea mays* ssp. *parviglumis* or spp. *mexicana*) (Beadle 1939; Galinat 1983; Iltis 1983), and introduced into Africa around 1550 (Miracle 1965). Today, maize is an important staple cereal crop across sub-Saharan Africa and is also often used as feed for livestock (Kiruwa et al. 2016; Mahuku et al. 2015a). This energy dense crop (365 kcal/100 g) is comparable to wheat and rice, and contains roughly 72% starch, 10% protein, and 4% fat (Inglett 1970). Maize can be used to produce a variety of foods (Fig. 1.1) as well as non-consumables such as paper, paint, textiles, and medicine (DAFF 2017). Since 2010, maize has even been used in the production of bio-fuels, especially in the Unites States where it accounts for about 40% of all maize produced (Ranum et al. 2014).

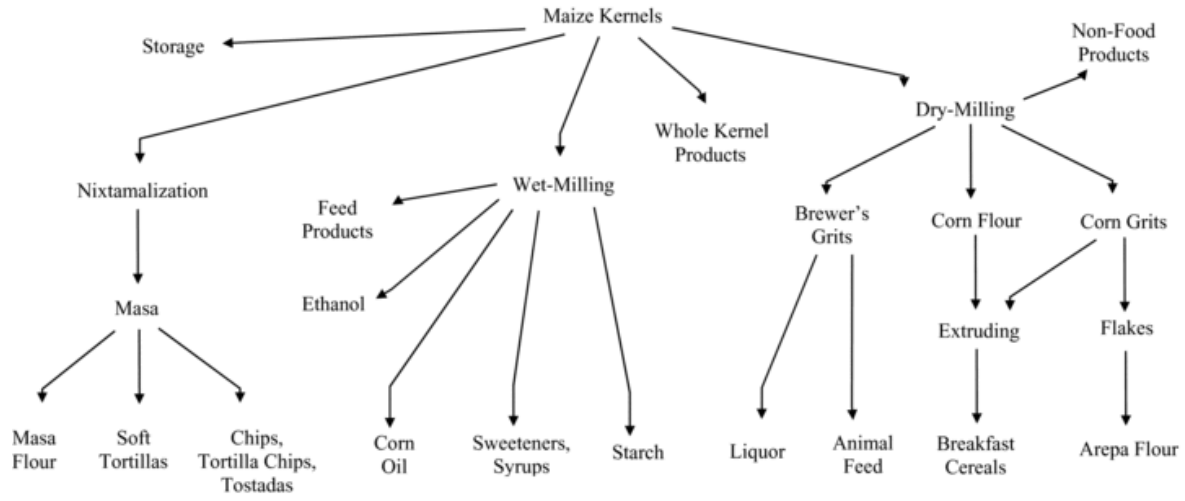


Fig. 1.1 The major processes and end products involved in raw kernel processing (Image reproduced from Nuss and Tanumihardjo 2010).

South Africa is Africa's largest maize producer, yielding an average of between 12 to 13 million tonnes annually (FAO 2018), with 2020's predicted yield well above the average at 16.1 million tonnes (FAO 2020). Maize is the largest produced crop in South Africa (Stats SA 2020), contributing approximately 15% of the gross value of all agricultural products (van Zyl and Nel 1988). This crop is cultivated over seven of South Africa's nine provinces (Free State, Mpumalanga, North West, Gauteng, KwaZulu-Natal, Limpopo and Northern Cape) on approximately 2.5 million hectares of land (FAO 2018). Of the total area planted, 87.5% is owned by commercial farmers which produce 94.6% of the total maize crop annually (Greyling and Pardey 2019).

The primary abiotic factors affecting maize production are drought, salinity, nutrient deficiencies, and high and low temperatures. Due to its heavy reliance on rainfall, maize is usually planted during the rainy seasons: October for the eastern regions of South Africa and between November and December for the western regions (FAO 2020). A study conducted by Adisa et al. (2018) looked into the effect of climate change on maize

production in South Africa over the period 1986–2015, taking precipitation, potential evapotranspiration, and minimum and maximum temperatures into account. Highest yields per hectare were recorded for the humid sub-tropical regions of KwaZulu-Natal and Mpumalanga, while the semi-arid regions of the Free State and North West produced the lowest yields per hectare (Adisa et al. 2018). The results showed that during the period of the study, the maximum temperature in all provinces increased and precipitation levels in the North West and Free State provinces decreased. The authors also predicted that these trends would likely continue in future. Despite this, the Free State still produces by far the most maize in South Africa, followed by Mpumalanga, and North West (Galal 2020).

In addition to abiotic stresses, biotic stresses too affect the maize industry with global yield loses of approximately 10% reported each year (Gong et al. 2014). One of the primary biotic stresses is that of viral pathogens, with over 50 viruses detected as naturally infecting maize, with maize identified as an experimental host for around 30 additional viruses (Lapierre and Signoret 2004). Of these, about 25 have been reported as causing economically significant yield losses (Lapierre and Signoret 2004).

1.2 Major viral diseases affecting maize in Africa

Maize streak disease (MSD) is considered the most economically important viral disease affecting maize in Africa (Rybicki 2015). Symptoms of MSD include white, yellow, or even red lesions on leaves, continuous parallel chlorotic streaks (Fig. 1.2A), plant stunting, and incomplete cob and seed formation resulting in severe yield losses (Shepherd et al. 2010). This disease is primarily caused by maize streak virus (MSV; genus: *Mastrevirus*, family *Geminiviridae*), which has been reported in many African countries and some Asian countries (Fig. 1.2B), and as widespread in South Africa since the 1870s (Fuller 1901).

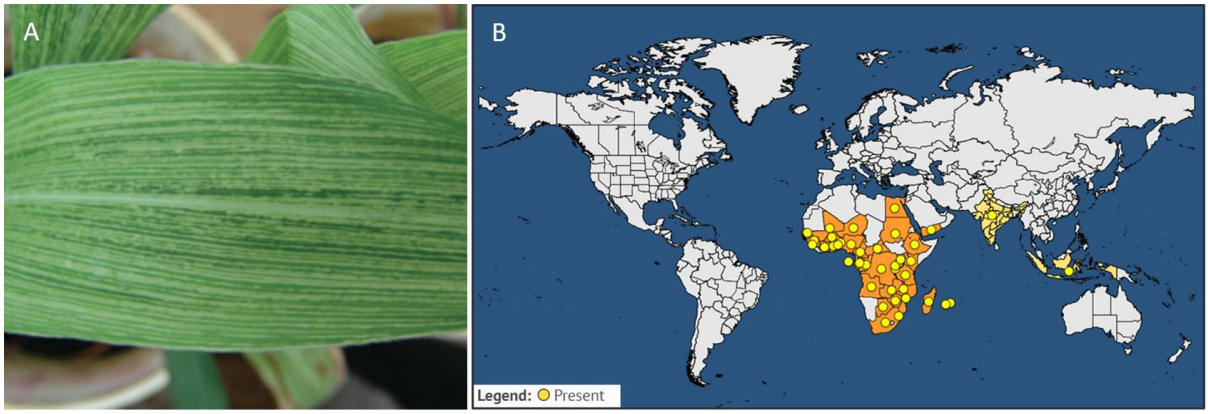


Fig. 1.2 Maize streak disease **(A)** chlorotic streaks on maize leaf caused by maize streak disease (Image reproduced from Shepherd et al. 2010); and **(B)** geographic distribution of maize streak virus in Africa and Asia (Image reproduced from EPPO Global Database 2019).

In 2011, a viral disease new to Africa, known as maize lethal necrosis disease (MLND), also known as corn lethal necrosis disease (CLND), was reported in the Southern Rift Valley of Kenya when a large outbreak occurred (Wangai et al. 2012). In this case, MLND was caused by the co-infection of maize with maize chlorotic mottle virus (MCMV; genus *Machlomovirus*; family *Tombusviridae*) and sugarcane mosaic virus (SCMV; genus

Potyvirus) (Wangai et al. 2012). However, co-infection of any maize-infecting member of the family *Potyviridae* with MCMV has been reported to cause MLND (Niblett and Claflin 1978).

Since the first report of MCMV in Peru in the early 1970s (Castillo and Hebert 1974), MCMV appeared to spread slowly, from South to North America, across to Thailand and Hawaii (Fig. 1.3) (Jiang et al. 1992; Niblett and Claflin 1978; Redinbaugh and Stewart 2018; Uyemoto 1983). However, from 2011, rapid emergence of the virus across southern Asia, sub-Saharan Africa and upwards into Spain was reported (Fig. 1.3) (De Groote et al. 2016; Deng et al. 2014; Kagoda et al. 2016; Mahuku et al. 2015a; Quito-Avila et al. 2016; Wangai et al. 2012; Xie et al. 2011). The spread of MLND to southern Tanzania (Read, et al. 2019a) may indicate that the disease is spreading southwards.

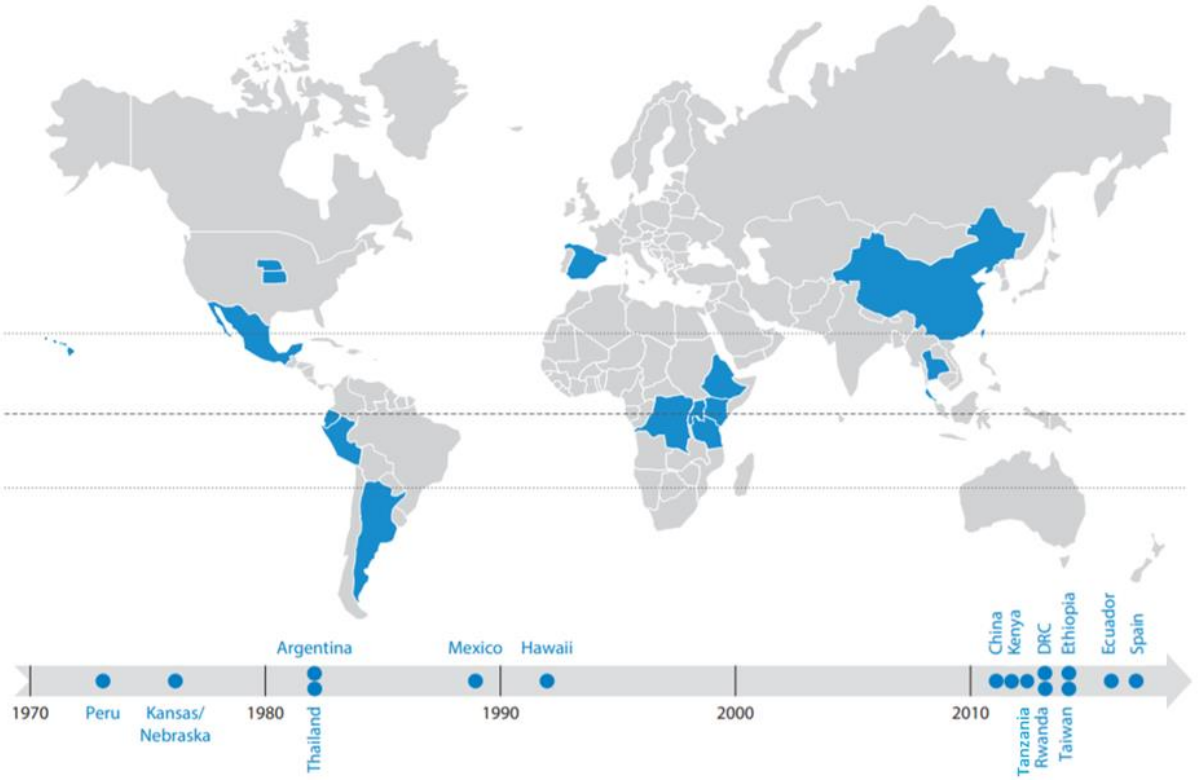


Fig. 1.3 Emergence of maize chlorotic mottle virus (MCMV). MCMV has been reported in a number of countries (blue), and within Africa primarily in Kenya, Rwanda, Democratic Republic of the Congo (DRC), Ethiopia and Tanzania. The reported year of MCMV emergence is indicated on the timeline (Image adapted from Redinbaugh and Stewart 2018b).

Within sub-Saharan Africa, reports of MCMV have been confirmed in Tanzania (FAO REOA 2013; Read et al. 2019c), Rwanda (Adams et al. 2014), the Democratic Republic of the Congo (Lukanda et al. 2014), Ethiopia (Mahuku et al. 2015b), and Uganda (Kagoda et al. 2016) (Fig.1.3), with suspected presence in South Sudan also reported but not confirmed (FAO REOA 2013). Crop losses of up to 50% were reported from individual farmers in Kenya and Uganda, with the maize yield losses to smallholder farmers in the affected African countries amounting to between 291 and 339 million USD for 2017 alone (Pratt et al. 2017). Thus, the presence of MLND in sub-Saharan Africa poses a serious threat to food security, where maize is used primarily as a staple food source (Mahuku et al. 2015a).

1.2.1 Primary viruses associated with MLND

1.2.1.1 Maize chlorotic mottle virus

MCMV seems to be the primary virus responsible for the spread of MLND, as global distribution is already described for potyvirids, especially SCMV, without the emergence of MLND (Redinbaugh and Stewart 2018). MCMV is currently the sole species of the genus *Machlomovirus* (King et al. 2011). It is a spherical, non-enveloped, 30 nm, icosahedral virus with a monopartite, 4.4 kb, linear, positive-sense, single-stranded RNA genome lacking a cap structure and a poly-A-tail (Scheets 2000).

The genome encodes a coat protein (cp), two movement proteins (p7a and 7pb), two RNA-dependent RNA polymerases (p50 and p111), a unique protein required for efficient systemic infection (p31) and a unique protein believed to play a role in virulence (p32) (Fig. 1.4) (Scheets 2000). Sub-genomic RNA1 (sgRNA1) expresses all the genes on the 3′ end of the genome as indicated in Fig. 1.4. Although further research is required for p31, p32, and p50 in order to further characterise the life cycle of this virus, there have been reports that these may be involved in host defence evasion via the suppression of RNA silencing (Csorba et al. 2015; Scheets 2016; Stenger and French 2008).

Current genome sequences available for different MCMV isolates suggest that the variants detected are very similar with only 1-4% nucleotide sequence diversity observed, but isolates from Africa and Asia appear more similar to each other than to other isolates (Mahuku et al. 2015a). This may potentially support the route of MCMV introduction into Africa through Asia as suggested by the timeline in Fig 1.3.

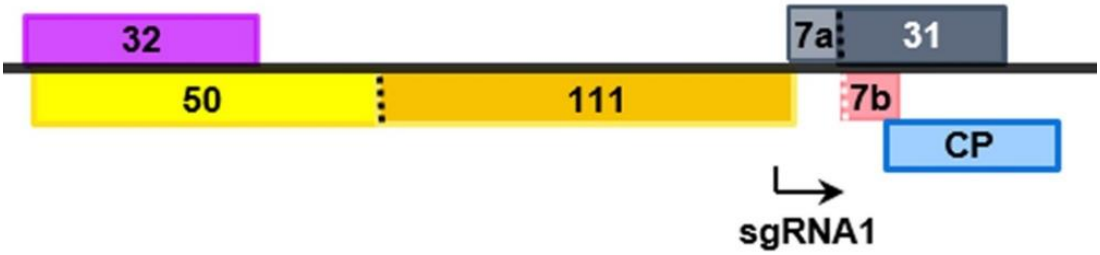


Fig. 1.4 Genome organization of maize chlorotic mottle virus, genus *Machlomovirus*, with abbreviations explained in text (Image reproduced from Scheets 2016).

Once MCMV is introduced into a host cell, the viral protein coat is removed and the genetic material released (Kiruwa et al. 2016). The RNA genome is first converted into complementary DNA (cDNA) using the host's machinery and enzymes, transcribed, and then translated into proteins that are used to produce multiple copies of the virus (Kiruwa et al. 2016). The newly synthesised viruses then spread to other cells through the plasmodesmata and then throughout the rest of the plant through the phloem tissue (Jeger et al. 2011). Thus, eventually the viral disease symptoms may be expressed systemically throughout the plant (Kiruwa et al. 2016).

1.2.1.2 Potyvirids

Other maize-infecting members of the family *Potyviridae* include maize dwarf mosaic virus (MDMV; genus *Potyvirus*), Johnsongrass mosaic virus (JGMV; genus *Potyvirus*) and wheat streak mosaic virus (WSMV; genus *Tritimovirus*) (Stewart et al. 2017). Co-infections with MDMV and WSMV were reported in Kansas (Niblett and Claflin 1978). Although potyvirid infections, especially SCMV, may have global distribution, to date, WSMV

has only been reported in the western hemisphere, eastern Europe, Australia and the Middle East, with no reports in East Africa or Asia (Hadi et al. 2011), and is not considered a major pathogen of maize (Redinbaugh and Stewart 2018).

Plant viruses in the family *Potyviridae* have non-enveloped, filamentous, flexuous particle structures with helical symmetry, and linear 8.2-11.3 kb positive-sense single-stranded RNA genomes with a 5′ genome-linked protein (VPg) and 3′ poly-A-tail (Wylie et al. 2017) with the general genome organization shown in Fig. 1.5. Potyviruses are usually monopartite (excluding *Bymovirus* spp.) with particles of 690-900 nm in length and 11-20 nm in diameter, while WSMV has a genome size of 9.4-9.6 kb with particles 15 nm in diameter (Stenger et al. 1998). Once introduced into the cells of the host plant, potyvirids experience a similar life cycle as described for MCMV as both have positive-sense, single-stranded RNA genomes. SCMV was reported as the most prominent potyvirus in maize in South Africa along with JGMV mainly infecting Johnsongrass and sweetcorn at low incidences (Schulze 2018).

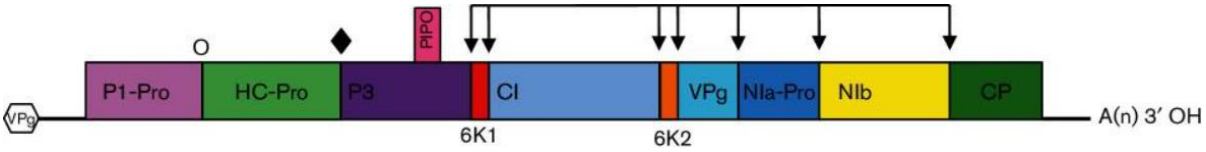


Fig. 1.5 Genome organization of a typical member of the genus *Potyvirus*. VPg, viral protein genome-linked; P1-Pro, protein 1 protease; HC-Pro, helper component protease; P3, protein 3; PIPO, pretty interesting *Potyviridae* open reading frame; 6K, six kilodalton peptide; CI, cytoplasmic inclusion; NIa-Pro, nuclear inclusion A protease; NIb, nuclear inclusion B body; RNA-dependent RNA polymerase; CP, coat protein. Cleavage sites of P1-Pro (O), HC-Pro (♦) and NIa-Pro (↓) are indicated (Image reproduced from Wylie et al. 2017).

1.2.2 Synergistic interaction between MCMV and potyvirids

Single infections of MCMV in maize may elicit symptoms such as stunting, chlorosis and mosaic (Fig. 1.6A). However, symptom severity has been noted to differ depending on the maize genotype, environmental conditions and time of infection (Mahuku et al. 2015a). Under unfavourable environmental conditions such as drought and low nitrogen availability, single infections of MCMV have been reported to result in symptoms similar to those of MLND (Flett and Mashingaidze 2016). Single infections of potyvirids, may also elicit very similar symptoms to those of MCMV (Fig. 1.6B) (Redinbaugh and Stewart 2018).

Co-infection of these viruses, on the other hand, elicits more severe symptoms such as chlorotic mottling on leaves, stunted growth, leaf necrosis, dead heart, small deformed ears with little to no seed set and premature death (Redinbaugh and Stewart 2018; Wangai et al. 2012) (Fig. 1.6C). The extent of the symptoms resulting from a mixed infection of MCMV and a potyvirid indicates that a synergistic interaction exists between the viruses rather than what would be expected from an additive effect. Co-infection of maize with a potyvirid has been reported to increase both MCMV and WSMV titres and siRNAs (Stenger et al. 2007), however, no effect on MDMV, SCMV and JGMV titres have been reported (Goldberg 1987; Stewart et al. 2017).

The HC-Pro, P1-Pro and nuclear inclusion proteins, NIa-Pro and NIb, of potyvirids are believed to aid the synergistic reaction with MCMV by interfering with RNA silencing of the host, thus allowing the viruses to evade the defence system of the host, encouraging replication and accumulation of both MCMV and the potyvirid, in the case of WSMV (Goldberg 1987; Mbega et al. 2016; Pruss et al. 1997; Wang et al. 2017).

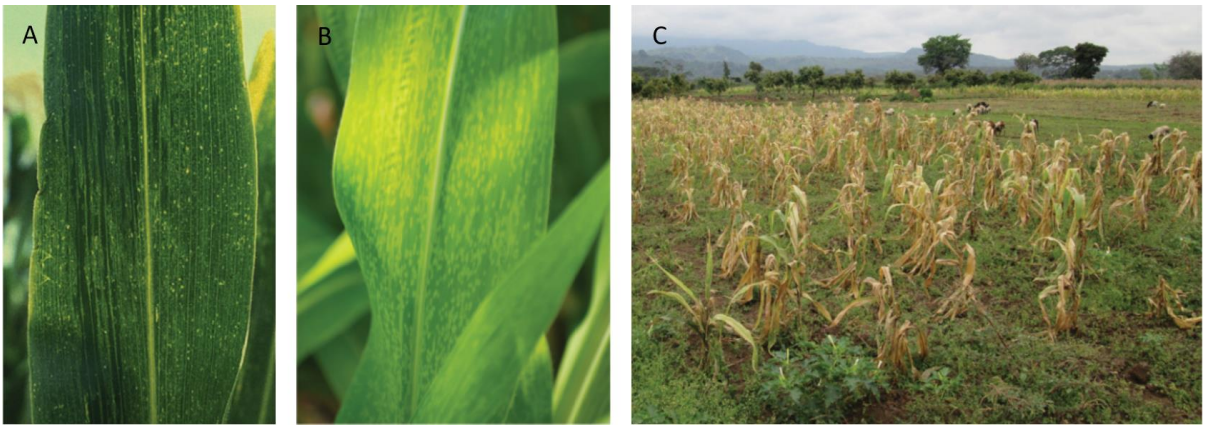


Fig. 1.6 Individual infection of **(A)** maize chlorotic mottle virus (MCMV); and **(B)** sugarcane mosaic virus (SCMV); followed by **(C)** a field with maize lethal necrosis disease resulting from co-infection of MCMV and SCMV (Images reproduced from Redinbaugh and Stewart 2018).

1.3 Additional MLND implicated viruses

Plants from Tanzania expressing MLND symptoms also contained infections of additional viruses such as MSV (Read et al. 2019a) and maize yellow mosaic virus (MaYMV; genus *Polerovirus*, family *Luteoviridae*) (also called maize yellow dwarf virus-RMV; MaYDV-RMV) (Read et al. 2019d), and recently described Morogoro maize-associated virus (MMaV; genus *Nucleorhabdovirus*, family *Rhabdoviridae*) (Read et al. 2019b), maize-associated pteridovirus (MaPV; genus *Pteridovirus*, family *Mayoviridae*) (Read et al. 2019c) and two maize-associated totivirus (MATV; unclassified genus in the family *Totiviridae*) variants, MATV-1-Tanz and MATV-4-Tanz (Read et al. 2019a). Possible symptoms and the effect on maize yield caused by MMaV, MaPV and MATV are yet to be determined. The roles of these viruses, including MSV, in mixed infection with MCMV are currently unknown.

1.3.1 Maize streak virus

MSV comprises a monopartite, 2.7 kb, circular, single-stranded DNA genome encapsidated in a 22 x 38 nm geminate structure with twinned incomplete icosahedral symmetry (Harrison et al. 1977). Like other grass-infecting mastreviruses, its genome comprises genes that encode three proteins, namely the capsid, movement and replication-associated proteins, as well as two untranslated regions known as the long and short intergenic regions (Zerbini et al. 2017) (Fig. 1.7). The conserved replication origin (5′-TAATATTAC-3′) is located within the long intergenic region and allows for the replicase-associated protein to cleave the sense strand, double stranded DNA to be formed using DNA polymerases of its host and subsequent bi-directional amplification to occur via rolling circle amplification (Hanley-Bowdoin et al. 2013).

Fig. 1.7 Genome organization of maize streak virus (MSV), genus *Mastrevirus*. The open reading frames (V1, V2, C1, and C2) are colour-coded according to the function of the protein products (rep, replication-associated protein; cp, capsid protein; mp, movement protein); LIR, long intergenic region; SIR, short intergenic region. The hairpin which includes the origin of replication is indicated in the LIR (Adapted from Varsani et al. 2014).

MSV is vectored by six leafhopper species in the genus *Cicadulina* (Storey 1924, 1925) and is known to infect a variety of both wild and cultivated grasses (Shepherd et al. 2010). Thus far, 11 types of MSV have been identified (MSV-A to -K) (Martin et al. 2001; Varsani et al. 2008). Only MSV-A is known to cause economically significant yield loss in maize, with MSV-B to -K predominantly infecting wild grass species (Shepherd et al. 2010) with only mild infections of MSV-B to -E reported in maize (Martin et al. 2001). The genetic variation observed has been attributed to the highly recombinant nature of MSV, with numerous intra-specific recombination events recorded for almost all types, except MVS-E , -G and -I (Monjane et al. 2011; Varsani et al. 2008). Initially, six subtypes of MSV-A were identified: MSV-A_1 to -A_6 (Martin et al. 2001), however, more recent studies reclassified MSV-A_5 as a group of recombinant variants that form a sublineage of the MSV-A_1 subtype (Owor et al. 2007), thus leaving five genetically distinct subtypes. Three MSV-B subtypes have also been identified (Varsani et al. 2008) with genomic sequences published on GenBank (National Center for Biotechnology Information, Bethesda, Maryland, USA) labelled based on the subtype they are believed to represent, MSV-B_1 to -B_3.

1.3.2 Maize yellow mosaic virus

Poleroviruses consist of a monopartite, 5-6 kb positive-sense single-stranded RNA genome with six open reading frames and three untranslated regions (Chen et al. 2016b). Poleroviruses are not mechanically transmissible but are vectored by the aphid *Rhopalosiphum maidis* (Chen et al. 2016b). The polerovirus, MaYMV, was first identified in China in 2016 (Chen et al. 2016b), and in 2018, MaYMV was detected in Pwani, Tanzania in a maize plant expressing MLND (Read et al. 2019d). Similar infections have also been reported in maize in Brazil (Gonçalves et al. 2017) and in sugarcane in Nigeria (Yahaya et al. 2017).

MaYMV infections have been reported as either asymptomatic or causing symptoms such as yellow mosaic (Chen et al. 2016b), and yellow streaking, possibly caused by co-infection with other viruses (Palanga et al. 2017; Welgemoed et al. 2020). Leaf reddening of MaYMV-infected plants has also been reported, however, these symptoms may be caused by the aphid vector, rather than the virus itself (Stewart et al. 2020). Alternate hosts of MaYMV include sugarcane, itchgrass (Yahaya et al. 2017) and sorghum (Lim et al. 2018; Wamaitha et al. 2018). It is currently unknown if this virus is soil or seed transmissible.

1.3.3 Morogoro maize-associated virus

Nucleorhabdoviruses are enveloped, 180 x 75 nm bullet-shaped particles that consist of a monopartite, linear, 11-15 kb negative-sense single-stranded RNA genome with five to six open reading frames (Jackson et al. 2005). A

recently discovered species of the genus *Nucleorhabdovirus*, tentatively named MMaV (Read et al. 2019b), has a genome length of about 12.2 kb with gene organization of 3′- nucleocapsid protein, phosphoprotein, movement protein, matrix protein, glycoprotein, and RNA-dependant RNA polymerase -5′, with nine terminal nucleotides on either end of the genome, displaying inverted complementarity as described for other rhabdoviruses (Read et al. 2019b). MMaV was reported in Tanzania in a maize plant expressing MLND, and is also the first report of a maize-infecting rhabdovirus in Africa (Read et al. 2019a). The etiology and epidemiology of this virus are currently unknown, however, since other plant nucleorhabdoviruses with monocot hosts are transmitted by leafhoppers and planthoppers (Whitfield et al. 2018), it is possible that MMaV may also be transmitted by one of these vectors.

1.3.4 Maize-associated pteridovirus

In February 2019, the pteridovirus, MaPV, was detected in Tanzania for the first time (Read et al. 2019c), and has since been found in Rwanda (Asiimwe et al. 2020) and South Sudan, according to a sequence record on GenBank (accession: MF372913). MaPV has a bipartite double-stranded RNA genome (Read et al. 2019c) with a 5.8 kb long RNA1 encoding a polyprotein product comprising of putative viral methyltrasferase, helicase and polymerase domains (Read et al. 2019c), and a 2.7 kb long RNA2 comprising three open reading frames encoding a putative movement protein, and two putative products of unknown function (Read et al. 2019b). Although the etiology of this virus is unknown, symptoms such as stunting, ringspot and necrosis have been associated with another genus member, the Japanese holly fern mottle virus, implying that species of the genus *Pteridovirus* may elicit severe disease symptoms (Read et al. 2019c; Valverde and Sabanadzovic 2009). The replication of MaPV is likely similar to that described for other double stranded RNA viruses, where once the virus enters the host cell, replication of the double-stranded RNA occurs inside the intact coat protein, preventing the host's immune system from being triggered (Liu and Cheng 2015).

1.3.5 Maize-associated totivirus

Totiviruses have been shown to infect a wide variety of fungi (Ghabrial et al. 2015), parasitic protozoa (Gómez-Arreaza et al. 2017), arthropods (Huang et al. 2018) and, most recently, plants including *Zea mays* (maize) (Alvarez-Quinto et al. 2017; Chen, Cao, et al. 2016; Read et al. 2019a). Totiviruses may have been introduced to plant hosts via fungal colonisation (Roossinck 2018). Totiviruses typically consist of a monopartite, 4.0-8.5 kb double-stranded RNA genome with two open reading frames (Read et al. 2019a). The 5′ open reading frame encodes a coat protein while the 3′ open reading frame encodes an RNA-dependent-RNA-polymerase (Read et al. 2019a).

Recently, maize-associated totiviruses have been reported in China (Chen et al. 2016a) and Ecuador (Alvarez-Quinto et al. 2017). Two MATV variants were also recently reported in Tanzania: MATV-1-Tanz has a genome length of 5,006 bp while MATV-4-Tanz (a divergent strain) has a genome length of 5,583 bp (Read et al. 2019a). Although it is possible that these viruses do infect the plant and are not merely present in fungi growing in/on the plant, all known attempts at elucidating the true host have been unsuccessful (Alvarez-Quinto et al. 2017; Chen et al. 2016a). Replication of MATV is also likely to occur within the intact coat protein once inside the host cell due to the double-stranded RNA nature of its genome (Liu and Cheng 2015).

1.4 Vectors

1.4.1 Maize chlorotic mottle virus vectors

In MLND symptomatic plants, MCMV occurs at high concentrations (Wang et al. 2017), making it possible for not only sucking insects such as thrips (*Frankliniella* spp.) (Cabanas et al. 2013), but also biting insects such as rootworms (*Diabrotica* spp.) (Jensen 1985; King et al. 2011), chrysomelid beetles (Nault et al. 1978) and stem borers (Mekureyaw 2017) to transmit this virus.

The most common MCMV vector in maize-growing regions of Africa is thrips (Kiruwa et al. 2016; Mahuku et al. 2015a, b). Sharma and Misra (2011) reported that after feeding on MCMV-infected maize for 3 h, thrips were able to infect other plants in a non-persistent manner, while Cabanas et al. (2013) suggested MCMV transmission by thrips occurred in a semi-persistent manner, with no latent period. Both Sharma and Misra (2011) and Cabanas et al. (2013) reported that after acquisition of the virus, thrips are able to transmit MCMV for up to six days. It is believed that longer feeding periods result in greater MCMV transmission efficiency, with the rate of transmission decreasing over time (Awata et al. 2019; Cabanas et al. 2013).

1.4.2 Potyvirid vectors

The most common vector of potyviruses (SCMV, JGMV and MDMV) are aphids, which transmit these viruses in a non-persistent manner (Brault et al. 2010), while WSMV is transmitted in a persistent manner by the eriophyid wheat curl mite (*Aceria tulipae* Keifer) (Somsen and Sill 1970). Over-wintering of aphids on infected alternate weed hosts was found to enhance the spread of potyvirids in the early stages of the following maize-growing season. Aphids, facilitated by wind turbulence, have also been reported to travel long distances between maize fields (Sharma and Misra 2011).

1.5 Virus reservoirs

MCMV has been reported in maize and other grasses such as sorghum, barley, wheat, millet, sugarcane, and weedy grasses, as well as a non-grass host, *Commelina benghalensis* (Bockelman 1982; Tonui et al. 2020; Wang et al. 2014) suggesting alternative hosts are likely important in the spread of this virus. Mechanical damage such as the use of unwashed utensils that have previously been used on infected tissue, have also been considered responsible for the spread of MCMV from plant to plant (Jensen et al. 1991; Kiruwa et al. 2016). Alternate hosts such as sugarcane, sorghum, cassava, beans, onion, rice and peppers, peas, coriander, Johnsongrass and other grass species have been reported as alternate hosts for potyvirids (Awata et al. 2019; Liu et al. 2017; Schulze 2018).

Seed transmission has been reported for both MCMV (Jensen et al. 1991) and potyviruses with transmission rates of between 0.2 and 0.4% reported for maize-infecting *Potyvirus* spp. (Shepherd and Holdeman 1965) and 0.1% for WSMV (Hill et al. 1974). Transmission of MCMV and potyvirids via infested soil, either mechanically or by a vector such as nematodes of fungi, has also been described (Bond and Pirone 1970; Nyvall 1999). However, due to the high prevalence of aphids, seed and soil transmission are not considered major routes of transmission (Shukla et al. 1994).

1.6 MLND disease detection

Before plant diseases can be effectively controlled, accurate methods of pathogen detection are first required. To date, a variety of different diagnostic methods have been used to identify the presence of viruses in maize, ranging from symptom-based, to serological and nucleic-acid based detection.

1.6.1 Symptom-based diagnostics

In plants, the presence of virus-like symptoms is often the first sign of a possible virus infection. The identification of these symptoms is important in order to perform management practices such as roguing. However, the identification of viral infections based on symptom expression can be difficult for a variety of reasons. In maize, some virus-infections may express very mild, inconclusive symptoms or appear asymptomatic due to the maize variety infected or the stage of infection (Kiruwa et al. 2016). Furthermore, many maize-infecting viruses express similar symptoms (Redinbaugh and Stewart 2018). Other factors such as pest damage, herbicide use, somatic mutation, harsh environmental conditions (drought and low nitrogen availability), and other microbial infections may cause virus-like symptoms (Redinbaugh and Stewart 2018). Therefore, although symptom observation may be useful as a form of preliminary screening, in order to make accurate diagnoses, other more reliable diagnostic tests are required.

1.6.2 Serological methods

Serological methods involve the detection of virus particles based on antigen-antibody reactions, such as the enzyme-linked immune-sorbent assays (ELISA) including double antibody sandwich ELISA (DAS-ELISA), triple antibody sandwich ELISA (TAS-ELISA) and direct antigen coating (DAC-ELISA) (Edwards and Cooper 1985). ELISAs are commonly used as routine plant virus diagnostics for large sample sizes, as they are affordable, quick, robust and simple to perform (Kiruwa et al. 2016). However, the production of high-quality antisera required for the sensitive and specific detection of viral antigens is expensive, time-intensive and requires virology expertise (Boonham et al. 2014). Furthermore, antisera often cannot correctly differentiate between closely related virus strains, with different phenotypes of the coat proteins (antisera targets) often conserved among genus members (Boonham et al. 2014). To date, ELISAs have been developed to identify the presence of the major MLND causing viruses, MCMV (Wu et al. 2013), SCMV (Thorat et al. 2015), MDMV (Zhang et al. 2010) and WSMV, as well as the MSD causative agent, MSV (Dekker et al. 1988).

1.6.3 Nucleic acid-based methods

Plant viruses may also be identified using nucleic acid-based methods such as polymerase chain reaction (PCR) and next-generation sequencing (NGS). PCR uses primers designed to bind to and amplify specific regions of a genetic sequence. The PCR products (amplicons) may be visualised on a gel, and directly sequenced (Kiruwa et al. 2016). PCR is widely employed as a diagnostic method due to its high sensitivity, specificity, versatility and speed (Ward et al. 2004). However, the downsides of using PCR include the high cost of reagents, the chance of false-negatives due to the non-uniform distribution of some viruses throughout the plant (Rao et al. 2006) and limitations of not being able to detect diverged virus variants due to the primer design being based off of known sequences only (Rahman et al. 2013). Nevertheless, PCR is considered the best method for reliable routine plant

virus diagnostics (Kiruwa et al. 2016). PCR has been used for the detection of MCMV (Stewart et al. 2014), potyvirids (Zheng et al. 2010), MSV (Willment et al. 2001), and five additional RNA viruses recently detected in MLND-affected maize in Tanzania (Read et al. 2019a, b, c, d).

NGS, on the other hand, is a new sequencing technology that generates sequence data for any genetic material present in a sample in a non-specific fashion (Adams et al. 2013). From the NGS data, sequences can be assembled and identified by comparing them to similar sequences on the GenBank database (Boonham et al. 2014). For this reason, NGS has been used not only as a diagnostic method, but also for population studies, as multiple viruses and virus strains can be detected in a given sample, and in the detection and characterisation of novel viruses (Boonham et al. 2014). However, NGS is not widely used as a routine diagnostic method, especially for large sample sizes, as it is very expensive (Kiruwa et al. 2016). NGS has been used to study the virus populations present in maize plants expressing MLND, and has resulted in the identification of novel viruses potentially implicated in MLND symptom severity, namely MaPV, MMaV and two MATV variants (Read et al. 2019a, b, c).

1.7 Disease management

1.7.1 MLND resistance/tolerance research

The most effective, economically and environmentally viable method of MLND control would be through the use of MLND resistant maize lines. In 2015, a study was done in Africa that tested 25,000 locally grown maize varieties for MLND resistance/tolerance (Gowda et al. 2015). The results suggested that 95% of the varieties tested were susceptible to MLND, which was not unexpected as the major genes/quantitative trait loci (QTLs) associated with virus resistance are not common in most maize varieties (Redinbaugh and Zambrano 2014).

Since the 2011 outbreaks in East Africa, some hybrids and inbred lines tolerant to MLND have been developed by the International Maize and Wheat Improvement Center (CIMMYT), Kenya Agricultural and Livestock Research Organization (KALRO) and other collaborating partners (Awata et al. 2019). Using genome-wide association studies (Gowda et al. 2015), three QTLs with associations to MLND resistance were identified, and are being used to improve existing high performance varieties. Beyene et al. (2017) also reported the development of three inbred maize lines that presented MLND resistance.

To date, genetic resistance to maize-infecting potyvirids such as MDMV, SCMV and WSMV, are some of the most well studied in maize (Adams et al. 2013; Liu et al. 2017; Rao et al. 2004; Thorat et al. 2015; Xie et al. 2011), with SCMV resistance described as a quantitively inherited trait (Xia et al. 1999). Scmv1, Scmv2 and Scmv3 were detected as dominant loci, each conferring protection against early infection, late infection, and throughout the life of the plant, respectively (Liu et al. 2017; Redinbaugh et al. 2004; Soldanova et al. 2012; Zhang-Ying et al. 2008).

Very little is known about MCMV tolerance. One study reported 47 out of 103 maize lines tested produced few to no symptoms, and were thus described as MCMV tolerant (Brewbaker and Martin 2015). Another study identified a QTL that reduced virus symptoms elicited in individual infections of both MCMV and SCMV (Gowda et al. 2018). There are currently no reports that describe the identification of maize lines with MCMV resistance.

1.7.2 Farming strategies

To give farmers the best chance of having a MLND-free maize growing season, the first important step is that the seed planted is MCMV-free. The best way to ensure this is for farmers to only sow certified seeds obtained from seed companies and avoid retaining seed produced from previous seasons (Awata et al. 2019). This may be more expensive initially but pales in comparison to the yield losses that may be incurred (between 291 and 339 million USD per year as mentioned in section 1.2) should a MLND outbreak occur.

Insecticides may be used to control vector abundance during the growing season (Jiang et al. 1992). Farmers should, however, be aware of the unintended negative effects such insecticides may have on ecological diversity (Quinn et al. 2011). The use of transgenic pest resistant maize may also prove useful in MLND vector control. For example, potential RNA interference (RNAi) targets have been identified against the MCMV vector *D. vigifera*, and other *Diabrotica* spp. (Nault et al. 1978). RNAi of V-ATPase-B of western flower thrips (*Frankliniella occidentalis*) has been found to successfully reduce gene expression and protein production, effectively reducing fertility and increasing mortality of females (Badillo-Vargas et al. 2015).

Plants identified as MCMV-infected should be removed immediately and infected grain and ears burned (Mawishe and Chacha 2013). MCMV-infected grain is unfit for human or animal consumption as secondary fungal infection associated with MLND may produce harmful aflatoxins (Farrar and Davis 1991; Parsons and Munkvold 2010; Redinbaugh and Stewart 2018; Yard et al. 2013). By removing and destroying infected plants, the spread of the disease can be reduced significantly (Kiruwa et al. 2016).

One of the reasons MLND is thought to have spread so quickly in East Africa may be due to the lack of rotational cropping employed, where maize is planted consecutively due to the presence of two rain seasons experienced in countries near the equator (De Groote et al. 2016; Fentahun et al. 2017). The major issues with continuous farming of maize, other than nutrient depletion, are that contaminated plant debris and soil accumulate, serving as a continual virus source (De Groote et al. 2016), and that the maize itself serves as a food source and breeding ground for virus vectors such as thrips, thus perpetuating the cycle (Redinbaugh and Stewart 2018).

Rather than planting maize continuously, crop rotation with legumes or other non-hosts is encouraged as it not only replenishes nitrogen in the soil (Havlin et al. 1990) but it also creates a break in the virus transmission cycle, thus reducing viral inoculum between maize growing seasons (Uyemoto 1983). Alternate hosts should be removed as these may facilitate virus persistence between seasons (Phillips et al. 1982; Uyemoto 1983) and farming machinery that may have been in contact with infected plants should be cleaned thoroughly (Kiruwa et al. 2016) to prevent mechanical transmission.

In addition to this, it is suggested that a policy be made by local authorities to enforce a minimum period of two months each year where no maize is permitted to be grown (Mezzalama et al. 2015). It is also suggested that weekly monitoring of fields be performed during these off-periods, whereby any volunteer maize plants be removed and destroyed (Mezzalama et al. 2015).

1.7.3 Authority-based management

Since potyvirids, especially SCMV, have a global distribution, outbreaks of MLND are usually caused by the introduction of MCMV (Redinbaugh and Stewart 2018). To prevent the spread of MCMV, phytosanitary regulations need to be put in place to prevent the importation of potentially infected plant material (including seed that may serve as a virus reservoir) into currently MLND-free countries (Mezzalama et al. 2015; Miano et al.

2013). Unfortunately, due to porous borders and inefficient implementation of such regulations in eastern and central Africa, MLND is expected to spread to all African countries neighbouring those already infected (Mekureyaw 2017).

Since smallholder famers make up the majority of maize farmers in sub-Saharan Africa, and since many of them may be poverty stricken, compliance with the suggested management practices recommended above may prove difficult. Many smallholders may not be able to afford the costs related to pesticide use, or purchasing certified seed at the start of each season, rather relying on seed retained from previous harvests (De Groote et al. 2016). Thus, in an attempt to manage the spread of the disease and reduce crop losses, governments of countries affected by MLND may consider supplying certified MLND-free seed and insecticides to farmers in need (Schulze 2018). Effort should also be made to educate farmers on suitable management strategies, some of which are mentioned above, to minimise virus transmission. Research into the development of MLND resistant/tolerant maize varieties should be prioritised and regular MLND virus surveillance encouraged, and funds for such research made available.

1.8 South Africa's predisposition to MLND

To date, MLND has not been reported in South Africa. This is believed to be because MCMV is still absent from the country. However, there have been reports of potyviruses in South Africa: MDMV was reported about 30 years ago (von Wechmar et al. 1987) and recently, albeit at low incidences, potyviruses such as SCMV and JGMV have been detected in maize grown in South Africa (Schulze 2018). The most recent reports of MSV distribution and diversity in South Africa, conducted on plants sampled around 20 years ago, suggested that MSV-A_1 and MSV-A_4 are wide-spread (Martin et al. 2001), with MaYMV also reported on maize earlier this year (Welgemoed et al. 2020). The virus status of MMaV, MaPV and MATV in South Africa is currently unknown.

Based on the potential southward expansion of MCMV through Africa predicted by Isabirye and Rwomushana (2016), it is anticipated that the virus may reach the sub-tropical areas of the country from Tanzania via Mozambique and/or Zimbabwe (Flett and Mashingaidze 2016). This is a serious concern as South Africa has already been flagged as a high-risk country with ideal humid sub-tropical conditions across large areas of South Africa's maize production regions, with KwaZulu-Natal, Mpumalanga and Limpopo being most at risk (Fig. 1.8) (Isabirye and Rwomushana 2016).

Furthermore, some of the major MCMV and potyvirid vectors, such as thrips (Allsopp 2010) and aphids (Hatting et al. 1999), are already known to occur in South Africa, and the climate of South Africa has been described as ideal for other MCMV vectors, such as the western corn rootworm, *Diabrotica virgifera virgifera*, to thrive should they be introduced into the country (Aragón et al. 2010; Kriticos et al. 2012). The presence of these vectors may facilitate the spread of MLND should an incursion of MCMV into South Africa occur.

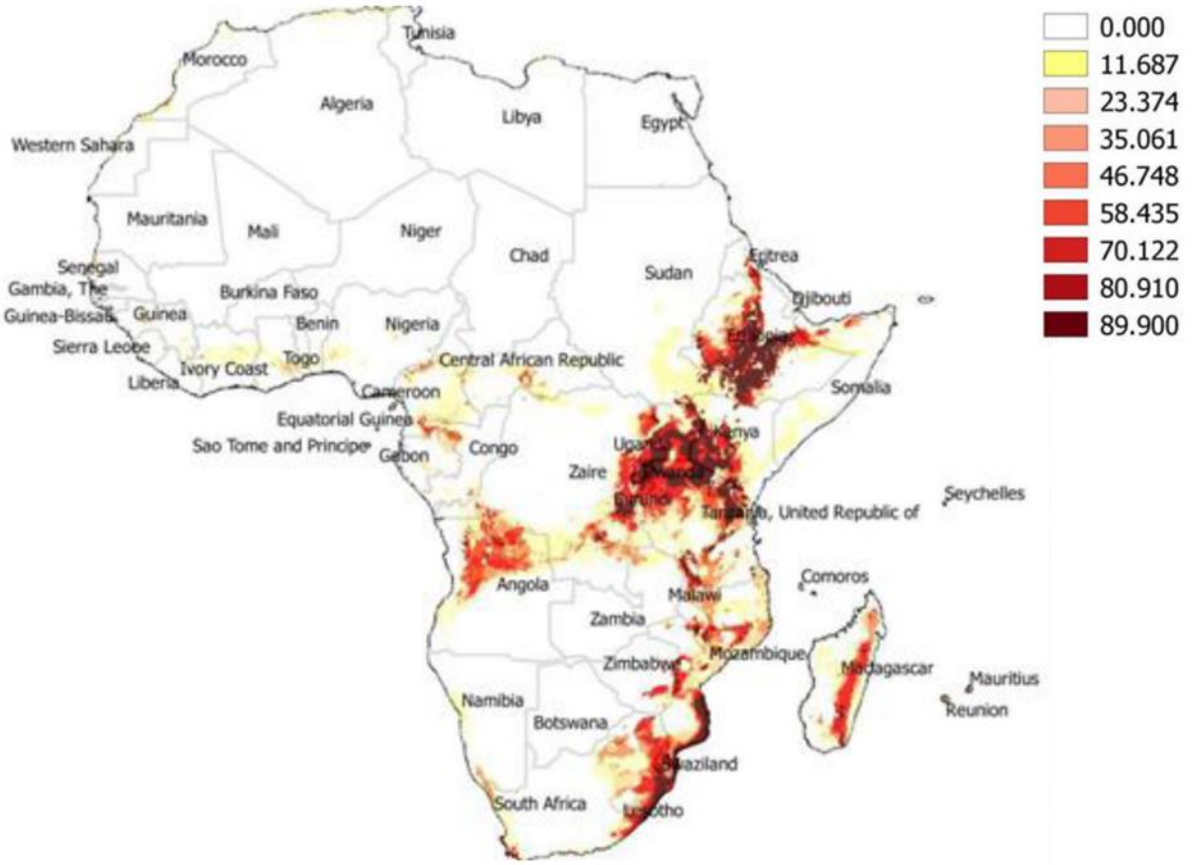


Fig. 1.8 Predicted potential distributions of maize chlorotic mottle virus and potential risk of maize lethal necrosis disease across Africa by 2050. Warmer colours indicate higher suitability and risk. (Image reproduced from Isabirye and Rwomushana 2016).

1.9 Conclusion

The spread of MLND across sub-Saharan Africa has resulted in dramatic crop losses. Research into this disease has revealed the synergistic reaction that takes place between MCMV and maize-infecting potyvirids. A variety of diagnostic systems have been developed for the detection of MLND-associated viruses, MLND and potyvirid resistant maize lines have been developed, and possible MCMV tolerant genotypes identified. The main insect vectors and modes of transmission of MLND-associated viruses have been determined and a variety of disease management strategies proposed.

Investigations into the nucleic acid extracted from maize plants showing MLND symptoms have also resulted in the identification of novel viruses including a nucleorhabdovirus, an array of polerovirus variants, a pteridovirus and an array of totivirus variants. Co-infections with MSV are also common. The effects that these viruses may have on maize and maize yield, the possible roles they may play in co-infections with MCMV and potyvirids, and their vectors and modes of transmission are all currently unknown.

In South Africa, potyviruses have already been reported, along with known virus vectors of both MCMV and potyviruses. The climatic conditions of the country are ideal for the survival of both MCMV and its vectors. Other viruses implicated in the MLND outbreak in East Africa, including MaYMV and MSV, have been detected in South Africa, although the most recent reports of MSV diversity and distribution were conducted on maize

plants sampled ~20 years ago. The presence and distribution in South Africa of other viruses such as MMaV, MaPV and variants of MATV remain unknown.

Future research efforts should be focussed on the development of MLND, MCMV and potyvirid resistant maize varieties, as well as on the etiology and epidemiology of the novel maize viruses detected, in order to determine the possible risk they may pose to the maize industry. Continued surveillance for MCMV should be conducted in countries neighbouring those where MLND has already been reported, and the implementation of effective phytosanitary regulations reinforced. To determine the risk of a MLND outbreak occurring in South Africa, pre-emptive surveillance of all MLND-implicated viruses should be performed.

1.10 References

Adams, I. P., Harju, V. A., Hodges, T., Hany, U., Skelton, A., Rai, S., et al. (2014). First report of maize lethal necrosis disease in Rwanda. *New Disease Reports*, 29, 22.

Adams, I. P., Miano, D. W., Kinyua, Z. M., Wangai, A., Kimani, E., Phiri, N., et al. (2013). Use of next-generation sequencing for the identification and characterization of maize chlorotic mottle virus and sugarcane mosaic virus causing maize lethal necrosis in Kenya. *Plant Pathology*. https://doi.org/10.1111/j.1365-3059.2012.02690.x

Adisa, O. M., Botai, C. M., Botai, J. O., Hassen, A., Darkey, D., Tesfamariam, E., et al. (2018). Analysis of agro-climatic parameters and their influence on maize production in South Africa. *Theoretical and Applied Climatology*, 134(3), 991–1004.

Aragón, P., Baselga, A., & Lobo, J. M. (2010). Global estimation of invasion risk zones for the western corn rootworm Diabrotica virgifera virgifera: integrating distribution models and physiological thresholds to assess climatic favourability. *Journal of Applied Ecology*, 47(5), 1026–1035.

Allsopp, E. (2010). Investigation into the apparent failure of chemical control for management of western flower thrips, Frankliniella occidentalis (Pergande), on plums in the Western Cape Province of South Africa. *Crop Protection*, 29(8), 824–831.

Alvarez-Quinto, R. A., Espinoza-Lozano, R. F., Mora-Pinargote, C. A., & Quito-Avila, D. F. (2017). Complete genome sequence of a variant of maize-associated totivirus from Ecuador. *Archives of Virology*, 162(4), 1083–1087.

Asiimwe, T., Stewart, L. R., Willie, K., Massawe, D. P., Kamatenesi, J., & Redinbaugh, M. G. (2020). Maize lethal necrosis viruses and other maize viruses in Rwanda. *Plant Pathology*, 69(3), 585–597.

Awata, L. A. O., Ifie, B. E., Tongoona, P., Danquah, E., Jumbo, M. B., Gowda, M., et al. (2019). Maize lethal necrosis and the molecular basis of variability in concentrations of the causal viruses in co-infected maize plant. *Journal of General and Molecular Virology*, 9(1), 1–19.

Badillo-Vargas, I. E., Rotenberg, D., Schneweis, B. A., & Whitfield, A. E. (2015). RNA interference tools for the western flower thrips, Frankliniella occidentalis. *Journal of Insect Physiology*, 76, 36–46.

Beadle, G. W. (1939). Teosinte and the origin of maize. *Journal of Heredity*, 30(6), 245–247.

Beyene, Y., Gowda, M., Suresh, L. M., Mugo, S., Olsen, M., Oikeh, S. O., et al. (2017). Genetic analysis of tropical maize inbred lines for resistance to maize lethal necrosis disease. *Euphytica*, 213(9), 224.

Bockelman, D. L. (1982). Host range and seed-transmission studies of maize chlorotic mottle virus in grasses and corn. *Plant Disease*. https://doi.org/10.1094/pd-66-216

Bond, W. P., & Pirone, T. P. (1970). Evidence for soil transmission of sugarcane mosaic virus. *Phytopathology*, 60(3), 437–440.

Boonham, N., Kreuze, J., Winter, S., van der Vlugt, R., Bergervoet, J., Tomlinson, J., & Mumford, R. (2014). Methods in virus diagnostics: From ELISA to next generation sequencing. *Virus Research*, 186, 20–31.

Brault, V., Uzest, M., Monsion, B., Jacquot, E., & Blanc, S. (2010). Aphids as transport devices for plant viruses. *Comptes Rendus Biologies*, 333(6–7), 524–538.

Brewbaker, J. L., & Martin, I. (2015). Breeding tropical vegetable corns. *Plant Breeding Reviews*, 39(39), 125–198.

Cabanas, D., Watanabe, S., Higashi, C. H. V, & Bressan, A. (2013). Dissecting the mode of maize chlorotic mottle virus transmission (Tombusviridae: Machlomovirus) by Frankliniella williamsi (Thysanoptera: Thripidae). *Journal of Economic Entomology*, 106(1), 16–24.

Castillo, J., & Hebert, T. T. (1974). A new virus disease of maize in Peru. *Fitopatologia*, 9(2), 79–84.

Chen, S., Cao, L., Huang, Q., Qian, Y., & Zhou, X. (2016a). The complete genome sequence of a novel maize-associated totivirus. *Archives of Virology*, 161(2), 487–490.

Chen, S., Jiang, G., Wu, J., Liu, Y., Qian, Y., & Zhou, X. (2016b). Characterization of a novel polerovirus infecting maize in China. *Viruses*, 8(5), 120.

Csorba, T., Kontra, L., & Burgyán, J. (2015). Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. *Virology*, 479, 85–103.

DAFF. (2017). A profile of the South African maize market value chain. *Commodity Profiles*. https://www.nda.agric.za/doaDev/sideMenu/Marketing/Annual Publications/Commodity Profiles/field crops/Maize Market Value Chain Profile 2017.pdf. Accessed 26 November 2020

De Groote, H., Oloo, F., Tongruksawattana, S., & Das, B. (2016). Community-survey based assessment of the geographic distribution and impact of maize lethal necrosis (MLN) disease in Kenya. *Crop Protection*, 82, 30–35.

Dekker, E. L., Pinner, M. S., Markkham, P. G., & Regenmortel, M. H. V. (1988). Characterization of maize streak virus isolates from different plant species by polyclonal and monoclonal antibodies. *Journal of General Virology*, 69, 983–990.

Deng, T. C., Chou, C. M., Chen, C. T., Tsai, C. H., & Lin, F. C. (2014). First report of maize chlorotic mottle virus on sweet corn in Taiwan. *Plant Disease*. https://doi.org/10.1094/PDIS-06-14-0568-PDN

Edwards, M. L., & Cooper, J. I. (1985). Plant virus detection using a new form of indirect ELISA. *Journal of Virological Methods*, 11(4), 309–319.

EPPO Global Database. (2019). Maize streak virus distribution. https://gd.eppo.int/taxon/MSV000/distribution. Accessed 25 November 2020

FAO. (2018). Country brief: South Africa. *GIEWS - Global Information and Early Warning System on Food and Agriculture*. http://www.fao.org/giews/countrybrief/country.jsp?code=ZAF. Accessed 13 March 2019

FAO. (2020). Country brief: South Africa. *GIEWS - Global Information and Early Warning System on Food and Agriculture*. http://www.fao.org/giews/countrybrief/country.jsp?code=ZAF&lang=ar. Accessed 25 November 2020

FAO REOA. (2013). Maize lethal necrosis disease (MLND) - A snapshot. *GIEWS - Global Information and Early Warning System on Food and Agriculture*. http://www.fao.org/fileadmin/user_upload/emergencies/docs/MLND Snapshot_FINAL.pdf. Accessed 25 November 2020

Farrar, J. J., & Davis, R. M. (1991). Relationships among ear morphology, western flower thrips, and Fusarium ear rot of corn. *Phytopathology*, 81(6), 661–666.

Fentahun, M., Feyissa, T., Abraham, A., & Kwak, H. R. (2017). Detection and characterization of maize chlorotic mottle virus and sugarcanemosaic virus associated with maize lethal necrosis disease in Ethiopia: an emerging threat to maize production in the region. *European Journal of Plant Pathology*, 149(4), 1011–1017.

Flett, B., & Mashingaidze, K. (2016). Maize lethal necrosis: Possible threat to local maize production. Grain SA. http://www.grainsa.co.za/maize-lethal-necrosis:-possible-threat-to-local-maize-production. Accessed 13 March 2020

Fuller, C. (1901). Mealic variegation. *First Report of the Government Entomologist Natal 1899-1900* (pp. 17–19). Pietermaritzburg: P. Davis & Sons, Government Printers.

Galal, S. (2020). Maize production in South Africa by province in 2018/2019. Statista. https://www.statista.com/statistics/1135488/maize-production-in-south-africa-by-province/. Accessed 26 November 2020

Galinat, W. C. (1983). The origin of maize as shown by key morphological traits of its ancestor, teosinte. Maydica.

Ghabrial, S. A., Castón, J. R., Jiang, D., Nibert, M. L., & Suzuki, N. (2015). 50-plus years of fungal viruses. *Virology*, 479, 356–368.

Goldberg, K.-B. (1987). Concentration of maize chlorotic mottle virus increased in mixed infections with Maize dwarf mosaic virus, strain B. *Phytopathology*. https://doi.org/10.1094/phyto-77-162

Gómez-Arreaza, A., Haenni, A.-L., Dunia, I., & Avilan, L. (2017). Viruses of parasites as actors in the parasite-host relationship: A "ménage à trois." *Acta tropica*, 166, 126–132.

Gonçalves, M. C., Godinho, M., Alves-Freitas, D. M. T., Varsani, A., & Ribeiro, S. G. (2017). First report of maize yellow mosaic virus infecting maize in Brazil. *Plant Disease*, 101(12), 2156.

Gong, F., Yang, L., Tai, F., Hu, X., & Wang, W. (2014). "Omics" of maize stress response for sustainable food production: opportunities and challenges. *Omics: A Journal of Integrative Biology*, 18(12), 714–732.

Gowda, M., Beyene, Y., Makumbi, D., Semagn, K., Olsen, M. S., Bright, J. M., et al. (2018). Discovery and validation of genomic regions associated with resistance to maize lethal necrosis in four biparental populations. *Molecular Breeding*, 38(5), 66.

Gowda, M., Das, B., Makumbi, D., Babu, R., Semagn, K., Mahuku, G., et al. (2015). Genome-wide association and genomic prediction of resistance to maize lethal necrosis disease in tropical maize germplasm. *Theoretical and Applied Genetics*. https://doi.org/10.1007/s00122-015-2559-0

Greyling, J. C., & Pardey, P. G. (2019). Measuring maize in South Africa: The shifting structure of production during the twentieth century, 1904–2015. *Agrekon*, 58(1), 21–41.

Hadi, B. A. R., Langham, M. A. C., Osborne, L., & Tilmon, K. J. (2011). Wheat streak mosaic virus on wheat: Biology and management. *Journal of Integrated Pest Management*, 2(1), J1–J5.

Hanley-Bowdoin, L., Bejarano, E. R., Robertson, D., & Mansoor, S. (2013). Geminiviruses: Masters at redirecting and reprogramming plant processes. *Nature reviews. Microbiology*, 11(11), 777–788.

Harrison, B. D., Barker, H., Bock, K. R., Guthrie, E. J., Meredith, G., & Atkinson, M. (1977). Plant viruses with circular single-stranded DNA. *Nature*, 270, 760–762.

Hatting, J. L., Humber, R. A., Poprawski, T. J., & Miller, R. M. (1999). A survey of fungal pathogens of aphids from South Africa, with special reference to cereal aphids. *Biological Control*, 16(1), 1–12.

Havlin, J. L., Kissel, D. E., Maddux, L. D., Claassen, M. M., & Long, J. H. (1990). Crop rotation and tillage effects on soil organic carbon and nitrogen. *Soil Science Society of America Journal*, 54(2), 448–452.

Hill, J. H., Martinson, C. A., & Russell, W. A. (1974). Seed transmission of maize dwarf mosaic and wheat streak mosaic viruses in maize and response of inbred lines 1. *Crop Science*, 14(2), 232–235.

Huang, Y., Guo, X., Zhang, S., Zhao, Q., Sun, Q., Zhou, H., et al. (2018). Discovery of two novel totiviruses from Culex tritaeniorhynchus classifiable in a distinct clade with arthropod-infecting viruses within the family Totiviridae. *Archives of Virology*, 163(10), 2899–2902.

Iltis, H. H. (1983). From teosinte to maize: The catastrophic sexual transmutation. *Science*, 222(4626), 886–894.

Inglett, G. E. (1970). Kernel structure, composition, and quality. In H.-U. Humpf (Ed.), *Molecular Nutrition and Food Research* (pp. 23-137). New Jersey: Wiley-Blackwell.

Isabirye, B. E., & Rwomushana, I. (2016). Current and future potential distribution of maize chlorotic mottle virus and risk of maize lethal necrosis disease in Africa. *Journal of Crop Protection*, 5(2), 215–228.

Jackson, A. O., Dietzgen, R. G., Goodin, M. M., Bragg, J. N., & Deng, M. (2005). Biology of plant rhabdoviruses. *Annual Review of Phytopathology*, 43, 623–660.

Jeger, M. J., van den Bosch, F., & Madden, L. V. (2011). Modelling virus- and host-limitation in vectored plant disease epidemics. *Virus Research*, 159(2), 215–222.

Jensen, S. G. (1985). laboratory transmission of maize chlorotic mottle virus by three species of corn rootworms. *Plant Disease*, 69(10), 864–868.

Jensen, S. G., Wysong, D. S., Ball, E. M., & Higley, P. M. (1991). Seed transmission of maize chlorotic mottle virus. *Plant Disease*, 75(5), 497–498.

Jiang, X. Q., Meinke, L. J., Wright, R. J., Wilkinson, D. R., & Campbell, J. E. (1992). Maize chlorotic mottle virus in Hawaiian-grown maize: Vector relations, host range and associated viruses. *Crop Protection*. https://doi.org/10.1016/0261-2194(92)90045-7

Kagoda, F., Gidoi, R., & Isabirye, B. E. (2016). Status of maize lethal necrosis in eastern Uganda. *African Journal of Agricultural Research*, 11(8), 652–660.

King, A. M. Q., Lefkowitz, E., Adams, M. J., & Carstens, E. B. (2011). *Virus taxonomy: Ninth report of the International Committee on Taxonomy of Viruses (Vol. 9).* San Diego: Elsevier.

Kiruwa, F. H., Feyissa, T., & Ndakidemi, P. A. (2016). Insights of maize lethal necrotic disease: A major constraint to maize production in East Africa. *African Journal of Microbiology Research*, 10(9), 271–279.

Kriticos, D. J., Reynaud, P., Baker, R. H. A., & Eyre, D. (2012). Estimating the global area of potential establishment for the western corn rootworm (Diabrotica virgifera virgifera) under rain-fed and irrigated agriculture. *EPPO Bulletin*, 42(1), 56–64.

Lapierre, H., & Signoret, P.-A. (2004). *Viruses and virus diseases of Poaceae (Gramineae)*. Paris: Editions Quae.

Lim, S., Yoon, Y., Jang, Y. W., Bae, D. H., Kim, B.-S., Maharjan, R., et al. (2018). First Report of Maize yellow mosaic virus Infecting Panicum miliaceum and Sorghum bicolor in South Korea. *Plant Disease*, 102(3), 689.

Liu, H., & Cheng, L. (2015). Cryo-EM shows the polymerase structures and a nonspooled genome within a dsRNA virus. *Science*, 349(6254), 1347–1350.

Liu, Q., Liu, H., Gong, Y., Tao, Y., Jiang, L., Zuo, W., et al. (2017). An atypical thioredoxin imparts early resistance to Sugarcane mosaic virus in maize. *Molecular Plant*, 10(3), 483–497.

Lukanda, M., Owati, A., Ogunsanya, P., Valimunzigha, K., Katsongo, K., Ndemere, H., & Kumar, P. L. (2014). First report of maize chlorotic mottle virus infecting maize in the Democratic Republic of the Congo. *Plant Disease*, 98(10), 1448.

Mahuku, G., Lockhart, B. E., Wanjala, B., Jones, M. W., Kimunye, J. N., Stewart, L. R., et al. (2015a). Maize lethal necrosis (MLN), an emerging threat to maize-based food security in sub-Saharan Africa. *Phytopathology*. https://doi.org/10.1094/PHYTO-12-14-0367-FI

Mahuku, G., Wangai, A., Sadessa, K., Teklewold, A., Wegary, D., Ayalneh, D., et al. (2015b). First report of Maize chlorotic mottle virus and maize lethal necrosis on maize in Ethiopia. *Plant Disease*, 99(12), 1870–1870.

Martin, D. P., Willment, J. A., Billharz, R., Velders, R., Odhiambo, B., Njuguna, J., et al. (2001). Sequence diversity and virulence in Zea mays of maize streak virus isolates. *Virology*. https://doi.org/10.1006/viro.2001.1075

Mawishe, R., & Chacha, E. (2013). *Uproot maize plants with lethal necrosis disease*. Wallingford: CABI

Mbega, E. R., Ndakidemi, P. A., Mamiro, D. P., Mushongi, A. A., Kitenge, K. M., & Ndomba, O. A. (2016). Role of potyviruses in synergistic interaction leading to maize lethal necrotic disease on maize. *International Journal of Current Microbiology and Applied Sciences*, 5(6), 85–96.

Mekureyaw, M. F. (2017). Maize lethal necrosis disease: An emerging problem for maize production in Eastern Africa. *Journal of Plant Physiology and Pathology*, 5(4), 1–6.

Mezzalama, M., Das, B., & Prasanna, B. M. (2015). MLN pathogen diagnosis, MLN-free seed production and safe exchange to non-endemic countries. Resource document. Research Program on Maize. https://doi.org/10.13140/RG.2.1.1234.3202

Miano, F., Kibaki, J., & Viollet, D. (2013). "Controlling maize lethal necrosis disease via vector management". *Fact sheet by Bayer Crop Science*.

Miracle, M. P. (1965). The introduction and spread of maize in Africa. *The Journal of African History*, 6(1), 39–55.

Monjane, A. L., Harkins, G. W., Martin, D. P., Lemey, P., Lefeuvre, P., Shepherd, D. N., et al. (2011). Reconstructing the history of maize streak virus strain a dispersal to reveal diversification hot spots and its origin in Southern Africa. *Journal of Virology*, 85(18), 9623–9636.

Nault, L. R., Styer, W. E., Coffey, M. E., Gordon, D. T., Negi, L. S., & Niblett, C. L. (1978). Transmission of maize chlorotic mottle virus by chrysomelid beetles. *Phytopathology*, 68, 1071–1074.

Niblett, C. L., & Claflin, L. E. (1978). Corn lethal necrosis - new virus disease of corn in Kansas. *The Plant Disease Reporter*, 62, 15–19.

Nuss, E. T., & Tanumihardjo, S. A. (2010). Maize: A paramount staple crop in the context of global nutrition. *Comprehensive Reviews in Food Science and Food Safety*, 9(4), 417–436.

Nyvall, R. F. (1999). *Field crop diseases*. Ames: Iowa State University Press.

Owor, B. E., Martin, D. P., Shepherd, D. N., Edema, R., Monjane, A. L., Rybicki, E. P., et al. (2007). Genetic analysis of maize streak virus isolates from Uganda reveals widespread distribution of a recombinant variant. *Journal of General Virology*. https://doi.org/10.1099/vir.0.83144-0

Palanga, E., Longué, R. D. S., Koala, M., Néya, J. B., Traoré, O., Martin, D. P., et al. (2017). First report of maize yellow mosaic virus infecting maize in Burkina Faso. *New Disease Reports*, 35, 26

Parsons, M. W., & Munkvold, G. P. (2010). Relationships of immature and adult thrips with silk-cut, fusarium ear rot and fumonisin B1 contamination of maize in California and Hawaii. *Plant Pathology*, 59(6), 1099–1106.

Phillips, N. J., Uyemoto, J. K., & Wilson, D. L. (1982). Maize chlorotic mottle virus and crop rotation: Effect of sorghum on virus incidence. *Plant Disease*, 66(5), 376–379.

Pratt, C. F., Constantine, K. L., & Murphy, S. T. (2017). Economic impacts of invasive alien species on African smallholder livelihoods. *Global Food Security*, 14(November 2016), 31–37.

Pruss, G., Ge, X., Shi, X. M., Carrington, J. C., & Vance, V. B. (1997). Plant viral synergism: The potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses. *The Plant Cell*, 9(6), 859–868.

Quinn, L., de Vos, J., Fernandes-Whaley, M., Roos, C., Bouwman, H., Kylin, H., et al. (2011). Pesticide use in South Africa: One of the largest importers of pesticides in Africa. In M. S. Rijeka (Ed.), *Pesticides in the Modern World: Pesticides Use and Management*. (pp. 49-96). Croatia: InTech

Quito-Avila, D. F., Alvarez, R. A., & Mendoza, A. A. (2016). Occurrence of maize lethal necrosis in Ecuador: A disease without boundaries? *European Journal of Plant Pathology*. https://doi.org/10.1007/s10658-016-0943-5

Rahman, M. T., Uddin, M. S., Sultana, R., Moue, A., & Setu, M. (2013). Polymerase chain reaction (PCR): A short review. *Anwer Khan Modern Medical College Journal*, 4(1), 30–36.

Ranum, P., Peña-Rosas, J. P., & Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. *Annals of the New York academy of Sciences*, 1312(1), 105–112.

Rao, G. P., Singh, M., Gaur, R. K., & Jain, R. K. (2004). Antigenic and biological diversity among sugarcane mosaic isolates from different geographical regions in India. *Indian Journal of Biotechnology*, 3(4), 538–541.

Rao, J. R., Fleming, C. C., & Moore, J. E. (2006). *Molecular diagnostics: Current technology and applications*. Horizon Bioscience Wymondham, Norfolk, UK.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019a). Diversity and distribution of maize-associated totivirus strains from Tanzania. *Virus G*enes, 55, 429–432.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019b). Molecular characterization of Morogoro maize-associated virus, a nucleorhabdovirus detected in maize (Zea mays) in Tanzania. *Archives of Virology*, 164, 1711–1715.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019c). Characterization and detection of maize-associated pteridovirus (MaPV), infecting maize (Zea mays) in the Arusha region of Tanzania. *European Journal of Plant Pathology*, 154, 1165–1170.

Read, D. A., Featherstone, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019d). First report of maize yellow mosaic virus (MaYMV) on maize (Zea mays) in Tanzania. *Journal of Plant Pathology*, 101(1), 203.

Redinbaugh, M. G., Jones, M. W., & Gingery, R. E. (2004). The genetics of virus resistance in maize (Zea mays L.). *Maydica*, 49(3), 183.

Redinbaugh, M. G., & Stewart, L. R. (2018). Maize lethal necrosis: An emerging, synergistic viral disease. *Annual Review of Virology*, 5, 301–322.

Redinbaugh, M. G., & Zambrano, J. L. (2014). Control of virus diseases in maize. In *Advances in Virus Research*. https://doi.org/10.1016/B978-0-12-801246-8.00008-1

Roossinck, M. J. (2018). Evolutionary and ecological links between plant and fungal viruses. *New Phytologist*, 221(1), 86–92.

Rybicki, E. P. (2015). A top ten list for economically important plant viruses. *Archives of Virology*, 160(1), 17–20.

Scheets, K. (2000). Maize chlorotic mottle machlomovirus expresses its coat protein from a 1.47-kb subgenomic RNA and makes a 0.34-kb subgenomic RNA. *Virology*, 267(1), 90–101.

Scheets, K. (2016). Analysis of gene functions in maize chlorotic mottle virus. *Virus Research*, 222, 71–79.

Schulze, S. E. (2018). *Pre-empting maize lethal necrosis disease in South Africa: Potyviruses of maize (Zea mays). M.Sc. Thesis*. University of Pretoria.

Sharma, K., & Misra, R. S. (2011). Molecular approaches towards analyzing the viruses infecting maize (Zea mays L.). *Journal of General and Molecular Virology*, 3(1), 1–17.

Shepherd, D. N., Martin, D. P., van der Walt, E., Dent, K., Varsani, A., & Rybicki, E. P. (2010). Maize streak virus: An old and complex "emerging" pathogen. *Molecular Plant Pathology*, 11(1), 1–12.

Shepherd, R. J., & Holdeman, Q. L. (1965). Seed transmission of the Johnson grass strain of the sugarcane mosaic virus in corn. *Plant Disease Reporter*, 49(6), 468–469.

Shukla, D. D., Ward, C. W., & Brunt, A. A. (1994). *The Potyviridae Wallingford*. UK: CAB International.

Smith, B. D. (1995). *The emergency of agriculture*. New York: Freeman.

Soldanova, M., Cholastova, T., Polakova, M., Piakova, Z., & Hajkova, P. (2012). Molecular mapping of quantitative trait loci (QTLs) determining resistance to sugarcane mosaic virus in maize using simple sequence repeat (SSR) markers. *African Journal of Biotechnology*, 11(15), 3496–3501.

Somsen, H. W., & Sill, W. H. (1970). *The wheat curl mite, Aceria tulipae Keifer, in relation to epidemiology and control of wheat streak mosaic*. Manhattan: Agricultural Experiment Station.

Stats SA. (2020). Census of commercial agriculture CoCA 2017. Resource Document. Stats SA http://www.statssa.gov.za/publications/Report-11-02-01/CoCA 2017 Fact Sheets.pdf

Stenger, D. C., & French, R. (2008). Complete nucleotide sequence of a maize chlorotic mottle virus isolate from Nebraska. *Archives of virology*, 153(5), 995–997.

Stenger, D. C., Hall, J. S., Choi, I., & French, R. (1998). Phylogenetic relationships within the family Potyviridae: Wheat streak mosaic virus and Brome streak mosaic virus are not members of the genus Rymovirus. *Phytopathology*, 88(8), 782–787.

Stenger, D. C., Young, B. A., Qu, F., Morris, T. J., & French, R. (2007). Wheat streak mosaic virus lacking helper component-proteinase is competent to produce disease synergism in double infections with maize chlorotic mottle virus. *Phytopathology*, 97(10), 1213–1221.

Stewart, L. R., Teplier, R., Todd, J. C., Jones, M. W., Cassone, B. J., Wijeratne, S., et al. (2014). Viruses in Maize and Johnsongrass in Southern Ohio. *Phytopathology*. https://doi.org/10.1094/PHYTO-08-13-0221-R

Stewart, L. R., Todd, J. C., Willie, K., Massawe, D. P., Khatri, & N. (2020). A recently discovered maize polerovirus causes leaf reddening symptoms in several maize genotypes and is transmitted by both the corn leaf aphid (Rhopalosiphum maidis) and the bird cherry-oat aphid (Rhopalosiphum padi). *Plant Disease*, 104(6), 1589–1592.

Stewart, L. R., Willie, K., Wijeratne, S., Redinbaugh, M. G., Massawe, D., Niblett, C. L., et al. (2017). Johnsongrass mosaic virus contributes to maize lethal necrosis in East Africa. *Plant Disease*. https://doi.org/10.1094/PDIS-01-17-0136-RE

Storey, H. H. (1924). The transmission of a new plant virus disease by insects. *Nature*, 114(2859), 245.

Storey, H. H. (1925). The transmission of streak disease of maize by the leafhopper Balclutha mbila Naude. *Annals of Applied Biology*, 12(4), 422–439.

Thorat, A. S., Pal, R. K., Shingote, P., Kharte, S., Nalavade, V., Dhumale, D. R., et al. (2015). Detection of sugarcane mosaic virus in diseased sugarcane using ELISA and RT-PCR technique. *Journal of Pure and Applied Microbiology*, 9(1), 319–327.

Tonui, R., Masanga, J., Kasili, R., Runo, S., & Alakonya, A. (2020). Identification of maize lethal necrosis disease causal viruses in maize and suspected alternative hosts through small RNA profiling. *Journal of Phytopathology*,168, 439–450.

Uyemoto, J. K. (1983). Biology and control. *Plant Disease*, 67(1), 7.

Valverde, R. A., & Sabanadzovic, S. (2009). A novel plant virus with unique properties infecting Japanese holly fern. *Journal of General Virology*, 90(10), 2542–2549.

van Zyl, J., & Nel, H. J. G. (1988). The role of the maize industry in the South African economy. *Agrekon*, 27(2), 10–16.

Varsani, A., Navas-Castillo, J., Moriones, E., Hernández-Zepeda, C., Idris, A., Brown, J. K., et al. (2014). Establishment of three new genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. *Archives of Virology*, 159(8), 2193–2203.

Varsani, A., Shepherd, D. N., Monjane, A. L., Owor, B. E., Erdmann, J. B., Rybicki, E. P., et al. (2008). Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. *Journal of General Virology*. https://doi.org/10.1099/vir.0.2008/003590-0

von Wechmar, M. B., Chauhan, R., Heam, S., & Knox, E. (1987). Applications of immunoelectroblotting to differentiate between strains of maize dwarf mosaic virus and sugarcane mosaic virus occurring in South Africa. In *CSFRI Symposium:(Research into Citrus and Subtropical Crops)* Abstract 59.

Wamaitha, M. J., Nigam, D., Maina, S., Stomeo, F., Wangai, A., Njuguna, J. N., et al. (2018). Metagenomic analysis of viruses associated with maize lethal necrosis in Kenya. *Virology Journal*. https://doi.org/10.1186/s12985-018-0999-2

Wang, Q., Zhang, C., Wang, C., Qian, Y., Li, Z., Hong, J., & Zhou, X. (2017). Further characterization of Maize chlorotic mottle virus and its synergistic interaction with sugarcane mosaic virus in maize. *Scientific Reports*, 7(1), 39960.

Wang, Q., Zhou, X. P., & Wu, J. X. (2014). First report of maize chlorotic mottle virus infecting sugarcane (Saccharum officinarum). *Plant Disease*, 98(4), 572.

Wang, R.-L., Stec, A., Hey, J., Lukens, L., & Doebley, J. (1999). The limits of selection during maize domestication. *Nature*, 398(6724), 236–239.

Wangai, A. W., Redinbaugh, M. G., Kinyua, Z. M., Miano, D. W., Leley, P. K., Kasina, M., et al. (2012). First report of maize chlorotic mottle virus and maize lethal necrosis in Kenya. *Plant Disease*, 96(10), 1582.

Ward, E., Foster, S. J., Fraaije, B. A., & McCartney, H. A. (2004). Plant pathogen diagnostics: Immunological and nucleic acid-based approaches. *Annals of Applied Biology*, 145(1), 1–16.

Welgemoed, T., Pierneef, R., Read, D. A., Schulze, S. E., Pietersen, G., & Berger, D. K. (2020). Next generation sequencing reveals past and current widespread occurrence of maize yellow mosaic virus in South Africa. *European Journal of Plant Pathology*, 158, 237–249.

Whitfield, A. E., Huot, O. B., Martin, K. M., Kondo, H., & Dietzgen, R. G. (2018). Plant rhabdoviruses - their origins and vector interactions. *Current Opinion in Virology*, 33, 198–207.

Willment, J. A., Martin, D. P., & Rybicki, E. P. (2001). Analysis of the diversity of African streak mastreviruses using PCR-generated RFLPs and partial sequence data. *Journal of Veterinary Medical Science*, 93, 75–87.

Wu, J., Wang, Q., Liu, H., Qian, Y., Xie, Y., & Zhou, X. (2013). Monoclonal antibody-based serological methods for maize chlorotic mottle virus detection in China. *Journal of Zhejiang University Science B*, 14(7), 555–562.

Wylie, S. J., Adams, M., Chalam, C., Kreuze, J., Jos, J., Ohshima, K., et al. (2017). ICTV Virus taxonomy profile: Potyviridae. *Journal of General Virology*, 98(3), 352–354.

Xia, X., Melchinger, A. E., Kuntze, L., & Lübberstedt, T. (1999). Quantitative trait loci mapping of resistance to sugarcane mosaic virus in maize. *Phytopathology*, 89(8), 660–667.

Xie, L., Zhang, J., Wang, Q., Meng, C., Hong, J., & Zhou, X. ping. (2011). Characterization of maize chlorotic mottle virus associated with maize lethal necrosis disease in China. *Journal of Phytopathology*. https://doi.org/10.1111/j.1439-0434.2010.01745.x

Yahaya, A., Al Rwahnih, M., Dangora, D. B., Gregg, L., Alegbejo, M. D., Lava Kumar, P., & Alabi, O. J. (2017). First report of maize yellow mosaic virus infecting sugarcane (saccharum spp.) and itch grass (rottboellia cochinchinensis) in Nigeria. *Plant Disease*, 101(7), 1335.

Yard, E. E., Daniel, J. H., Lewis, L. S., Rybak, M. E., Paliakov, E. M., Kim, A. A., et al. (2013). Human aflatoxin exposure in Kenya, 2007: A cross-sectional study. *Food Additives & Contaminants: Part A*, 30(7), 1322–1331.

Zerbini, F. M., Briddon, R. W., Idris, A., Martin, D. P., Moriones, E., Navas-Castillo, J., et al. (2017). ICTV virus taxonomy profile: Geminiviridae. *Journal of General Virology*, 98(2), 131–133.

Zhang-Ying, X. I., ZHANG, S.-H., Xin-Hai, L. I., Chuan-Xiao, X. I. E., Ming-Shun, L. I., Zhuan-Fang, H. A. O., et al. (2008). Identification and mapping of a novel sugarcane mosaic virus resistance gene in maize. *Acta Agronomica Sinica*, 34(9), 1494–1499.

Zhang, Z.-Y., Fu, F.-L., Gou, L., Wang, H.-G., & Li, W.-C. (2010). RNA interference-based transgenic maize resistant to maize dwarf mosaic virus. *Journal of Plant Biology*, 53(4), 297–305.

Zheng, L., Rodoni, B. C., Gibbs, M. J., & Gibbs, A. J. (2010). A novel pair of universal primers for the detection of potyviruses. *Plant Pathology*, 59, 211–220.

Chapter 2: Pre-empting maize lethal necrosis disease: Survey for viruses affecting maize in KwaZulu-Natal, South Africa.

2.1 Introduction

Maize lethal necrosis disease (MLND) causes maize crop losses for individual farmers of up to 100% (Mahuku et al. 2015a; Pratt et al. 2017; Redinbaugh and Stewart 2018), and is thus a of serious concern for food security, especially in sub-Saharan Africa where maize is the most important staple cereal crop (Kiruwa et al. 2016). MLND is caused by the synergistic co-infection of MCMV and a maize-infecting potyvirids, with the spread of MCMV recognised as the primary cause of MLND emergence since the potyvirids that are involved already have a global distribution (Redinbaugh and Stewart 2018). Since its emergence in Kenya in 2011 (Wangai et al. 2012), MLND has been confirmed in at least six African countries [Tanzania (FAO REOA 2013), Rwanda (Adams et al. 2014), the Democratic Republic of the Congo (Lukanda et al. 2014), Ethiopia (Mahuku et al. 2015b), and Uganda (Kagoda et al. 2016)], with recent reports of MLND in the southern parts of Tanzania (Read et al. 2019a). Concern exists that MCMV is spreading in a southward direction and is predicted to spread to South Africa via Mozambique and/or Zimbabwe (Isabirye and Rwomushana 2016).

South Africa is predisposed to MLND for a variety of reasons. Previous studies have already reported the presence of potyviruses such as SCMV, JGMV (Schulze 2018), and MDMV (von Wechmar et al. 1987) on maize in South Africa, along with major known vectors of MCMV (thrips) (Allsopp 2010) and potyviruses (aphids) (Hatting et al. 1999). South Africa's climate has been described as ideal for both MCMV, and its other known vectors, such as the western corn rootworm (Kriticos et al. 2012), to thrive in the country should they be introduced (Isabirye and Rwomushana 2016). Using a model based on climate suitability for the survival of MCMV and its vectors and accounting for predicted effects of climate change, KwaZulu-Natal, Limpopo and Mpumalanga are among the provinces most at risk of MLND outbreaks occurring should MCMV be introduced into the country (Isabirye and Rwomushana 2016).

The presence of other viruses were also detected in MLND-affected plants in Tanzania, including maize streak virus (MSV; a single-stranded DNA virus) (Read et al. 2019a), and five RNA viruses: maize yellow mosaic virus (MaYMV) (Read et al. 2019d), Morogoro maize associated virus (MMaV) (Read et al. 2019b), maize-associated pteridovirus (MaPV) (Read et al. 2019a) and two maize-associated totivirus (MATV) variants, MATV-1-Tanz and MATV-4-Tanz (Read et al. 2019c). Thus, pre-empting the introduction of MCMV into South Africa, the aim of the present study was to determine the current possible incidence and distribution of MCMV, potyvirids, MSV, MaPV, MMaV and two MATV variants in KwaZulu Natal, South Africa, the province predicted to be at highest risk of a MLND outbreak. The distribution of MaYMV, including in KwaZulu Natal, is currently the topic of a separate study (MSc. study of Ms M. Mamabola at Stellenbosch University).

2.2 Materials and methods

2.2.1 Sampling

In early January 2019, a survey was conducted along the major grain transport route leading between Gauteng and KwaZulu-Natal as well as the interior of KwaZulu-Natal as requested by the project funder, the South African National Seed Organization (SANSOR). Where available, commercial and smallholder maize fields were surveyed along the roadside approximately 20 km apart, while sites with only volunteer plants were sampled at a minimum of 5 km apart. Maize leaves with virus-like symptoms were collected from each site by four field technicians for between 20 to 40 person-minutes. Leaf symptoms were photographed, characteristics of each site recorded, and site co-ordinates saved on a Trimble GeoXH handheld differential global positioning system. For instances where many different symptoms or a high incidence of disease were observed, a maximum of 20 plants were sampled per site. Samples were placed in paper bags, dried with silica desiccant, and stored at room temperature until utilised.

2.2.2 Nucleic acid extraction and quality control

Dried maize leaf samples were pooled into groups of five and total nucleic acid extracted from ~200 mg (~40 mg per sample) of dry leaf tissue using a cetyltrimethylammonium bromide (CTAB) (2% CTAB, 1% PVP-40, 20 mM EDTA pH 8, 100 mM Tris-HCl pH 8, 1.4 M NaCl, and 3% β-mercaptoethanol) extraction method (White et al. 2008), modified by omitting spermidine from the CTAB buffer. The sample tissue and CTAB mixture were added to Universal 12 x 15 cm extraction bags (BIOREBA, Reinach, CH) and homogenised using a hand-held homogeniser (BIOREBA, Reinach, CH). The concentrations and purities of the extracts were analysed spectrophotometrically (NanoDrop® 1000; Thermo Scientific, Deleware, USA). Primers (Actin F: 5′- ACC GAA GCC CCT CTT AAC CC -3′ and Actin R: 5′- GTA TGG CTG ACA CCA TCA CC -3′) spanning an intron within the conserved Actin gene of plants (Inqaba Biotechnical Industries, Pretoria, ZA) were used in a one-step RT-PCR and served as a "housekeeping" control of the RNA (~180 bp amplicon) and DNA (~270 bp amplicon) for each extract (van den Berg et al. 2004). A OneTaq® One-Step RT-PCR kit (New England BioLabs, Massachusetts, USA) was utilised, as per the manufacturer's instructions, with the following reverse transcription PCR (RT-PCR) cycling conditions: reverse transcription at 48°C for 30 min, initial denaturation at 94°C for 1 min, followed by 40 cycles of amplification (denaturation at 94°C for 15 s; annealing at 42°C for 10 s; extension at 68°C for 15 s), followed by a final extension at 64°C for 5 min. The RT-PCR products were loaded onto a 1.5% TBE-agarose gel stained with ethidium bromide, and electrophoresed at 100 V for 50 min.

2.2.3 PCR and RT-PCR based virus diagnostics

A OneTaq® One-Step RT-PCR kit was used to test for the presence of MCMV, as per the manufacturer's instructions, using primers MCMV 894F and MCMV 1553R (Table 2.1). The RT-PCR cycling conditions were as follows: reverse transcription at 48°C for 30 min, initial denaturation at 94°C for 1 min, followed by 48 cycles of amplification (denaturation at 94°C for 15 s; annealing at 60°C for 30 s; extension at 68°C for 1 min), and a final extension at 68°C for 10 min. Reaction products were visualised by gel electrophoresis on a 1.5% agarose gel with TAE buffer at 100 V. Ethidium bromide staining and ultraviolet light transillumination was used to visualise the results.

Potyvirids, MMaV, MaPV, MATV-1-Tanz and MATV-4-Tanz were assayed using 1 µl nucleic acid in a one-step RT-PCR reaction mixture containing 0.5 µM of each respective forward and reverse primer (Table 2.1) 1 X GoTaq® buffer, 4 U Recombinant RNasin® ribonuclease inhibitor and 40 U Moloney Murine Leukemia Virus reverse transcriptase, 0.175 mM dNTP mix, 10 mM molecular grade DDT, 1.5 mM $MgCl_2$, and 0.5 U GoTaq® DNA polymerase, made to a total volume of 25 µl with nuclease-free water (all aforementioned reagents sourced from Promega Corporation, Madison, Wisconsin, USA). Reaction conditions were as follows: cDNA synthesis at 37°C for 45 min, elongation at 50°C for 2 min, initial denaturation at 94°C for 4 min followed by 40 amplification cycles (denaturation at 94°C for 30 s, annealing at 42°C for potyvirids/55°C for the other RNA viruses for 30 s, elongation at 72°C for 1 min), and a final extension at 72°C for 2 min. Amplification products were visualised as described previously, and bands of expected size purified using Zymoclean Gel DNA Recovery Columns (Zymo Research) and sent for bidirectional Sanger sequencing (Central Analytical Facility, DNA Sequencing Unit, Stellenbosch, ZA). Nucleic acid was extracted from those individuals of pools that tested positive for MMaV and MaPV, and the PCR's repeated on individual plant extracts of positive pools to determine the individual sources of infection.

Pooled maize extracts were tested for MSV using primers MSV F and MSV R (Table 2.1), and the same reaction mixture as above for the various RNA viruses, with the exclusion of 4 U Recombinant RNasin® ribonuclease inhibitor and 40 U Moloney Murine Leukemia Virus reverse transcriptase. PCR cycling conditions were as follows: initial template denaturation at 94°C for 4 min, followed by 40 amplification cycles (denaturation at 94°C for 30 s, annealing at 55°C for 30 s, extension at 72°C for 1 min 30 s) and a final extension at 72°C for 2 min. Reaction products were visualised, purified and sequenced as described above.

Table 2.1 Primers selected for maize virus detection.

Detection target	Primer name	Primer sequence (5′-3′)	Amplicon size (bp)	Reference
MCMV	MCMV 894F	TGGAACAGGCTATGGAACAGAATG	660	(Stewart et al. 2014)
	MCMV 1553R	TGCGGGTTTTGTGTCTCGTG		
Potyvirids	NIB 2F	GTITGYGTIGAYGAYTTYAAYAA	350	(Zheng et al. 2010)
	NIB 3R	TCIACIACIGTIGAIGGYTGNCC		
MSV	MSV F	CCAAAKDTCAGCTCCTCCG	~1,300	(Willment et al. 2001)
	MSV R	TTGGVCCGMVGATGTASAG		
MMaV	MMaV F	TCTGATTCTTGCCAAATGCTACC	766	(Read et al. 2019c)
	MMaV R	GATAATGAGCATCTCCACCAGAC		
MaPV	MaPV RNA2 F	CCGTTAACCGGAGATCCTACGA	687	(Read et al. 2019b)
	MaPV RNA2 R	CAGAGTACCAGCGACAGCATC		
MATV-1-Tanz	MATV 1F	CTACCTCCGATGCACAATGAGTTC	664	(Read et al. 2019a)
	MATV 1R	GGATAGAGTGCGCTTGACGATG		
MATV-4-Tanz	MATV 4F	ATCGTAGTGTGTCGTTCACAGG	763	(Read et al. 2019a)
	MATV 4R	CAACATTAGATCGTCTGCCGACG		

2.2.4 RNA-seq and bioinformatic analysis

RNA extracts were prepared of representative samples infected with viruses not previously reported in South Africa. These were submitted for next generation sequencing (NGS). A Qubit 3 instrument with broad range RNA detection reagent was used to determine the RNA concentrations and an Implen Nanophotometer (Implen, Munich, DE) to determine the A_{260}/A_{280} and A_{260}/A_{230} ratios. An RNAtag library was prepared according to

Shishkin et al. (2015) and sequenced using an Illumina HiSeq 2500 instrument (Illumina, San Diego, CA, US) (2 x 125 bp paired-end reads).

The read quality was analysed using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/) and trimmed using CLC Genomics Workbench 20.0.2 (https://digitalinsights.qiagen.com) with the following settings: read-through adapter sequences automatically removed using custom trim adaptor list; low-quality sequences removed (limit: 0.05); ambiguous nucleotides removed (maximum allowed: 2). Trimmed reads were then subjected to *de novo* assembly using a minimum contig length of 500 nt, length fraction of 0.9, and similarity fraction of 0.9. Contigs obtained were then subjected to BLASTn (National Center for Biotechnology Information, Bethesda, Maryland, USA) for analysis.

Thereafter, trimmed reads were mapped to relevant virus reference sequences having high similarity to the *de novo*-assembled contigs using the following parameters: no masking; match score: 1; mismatch cost: 2; insertion/deletion cost: 3; length fraction: 0.9; similarity fraction: 0.9; global alignment: yes; and non-specific match handling: map randomly (CLC). All consensus sequences obtained were analysed using BLASTn to determine the percentage identity to known GenBank variants. The consensus lengths and average coverage depths obtained were then recorded.

The *de novo*-assembled contigs and reference-mapped consensus sequences were then aligned using a gap open cost of 10 (or 100 where necessary) and a gap extension cost of 1. The ends of the contig sequences were trimmed to that of the reference sequence to remove base calls of low confidence. Each final draft genome/coding domain sequence (CDS) was then extracted, analysed with BLASTn and submitted to the GenBank database.

2.2.5 Additional MMaV confirmation

Two additional primer sets were designed with Primer3 (https://primer3.ut.ee/) using standard settings to confirm the presence of MMaV. The designed primers were as follows: 1) MMaV F3 (5′- GTG ATA GGG CAA AGA GGA GTC -3′) and MMaV R3 (5′- GCT GTA TGA AGA AGA TGG CTG G -3′) with an expected amplicon size of 481 bp; and 2) MMaV F4 (5′- GTG AAC TCT GCG TGG ACC TG -3′), and MMaV F4 (5′- GAG GGA RCR GAT GGA AGT CG -3′) with an expected amplicon size of 364 bp. The PCR and cycling conditions used were the same as those described in section 2.3 for the MMaV detection primers. The bands produced were then excised, sequenced, and identified using BLASTn.

2.2.6 Phylogenetic analysis

The final draft genome sequences/CDSs for MaPV and MMaV were transcribed into their amino acid (aa) sequences in CLC and aligned to the RNA-dependent RNA polymerase (RdRP) region of other virus references available on GenBank using MUSCLE alignment function in MEGA X (Kumar et al. 2018). The alignments were then trimmed to represent suitable cognate regions. A protein model test was then performed to determine the model that best fit each of the datasets and the respective models were then used to construct Maximum Likelihood phylogenies with 1,000 bootstrap replicates.

Using the MUSCLE alignment function, the complete genome sequences/CDSs of all known variants of MaPV (RNA1), MMaV, and MATV, along with suitable outgroups, were then aligned and trimmed to represent cognate regions. A nucleotide model test was then performed on each dataset and the model of best fit used to

construct Maximum Likelihood trees with 1,000 bootstrap replicates. The trees were then rooted with their respective outgroups.

2.3 Results

2.3.1 PCR and RT-PCR-based survey

A total of 50 sites were surveyed. Of these, virus-like symptoms were observed in 39 sites and 208 samples were collected (Fig. 2.1). Two asymptomatic samples were collected from two of the eleven sites where no symptoms were observed to serve as field grown healthy controls. No attempts were made to determine the maize cultivars collected. Virus-like symptoms observed on the leaves included chlorotic and necrotic lesions, interveinal chlorosis, chlorotic speckling, streaking, mild mottling, red vein banding, and reddening of leaf margins (Supplementary 2.1). Nucleic acid extracts made from a pool of five maize plants each had a minimum concentration of 200 ng/µl, an A_{260nm}/A_{280nm} greater than 2.0, A_{260nm}/A_{230nm} greater than 1.8, and contained both DNA and RNA, as confirmed by the actin RT-PCR. Virus-positive PCR amplicons from previous studies (Read et al. 2019a, b, c) were obtained to serve as positive controls.

RT-PCR diagnostic screening was performed on all 42 pools, of which 12, collected from nine different sites in Winterton, Mfekayi, Mkuze/Pongola, and central Pongola, tested positive for MSV (Table 2.2; Supplementary 2.1). Additionally, MaPV was also detected in one pool collected from Mfekayi, while another sampled from Pongola, also tested positive for MMaV (Table 2.2; Supplementary 2.1). No MCMV, potyvirids, MATV-1-Tanz or MATV-4-Tanz infections were detected (Table 2.2). RT-PCR assays of individual plants from virus-positive pools detected MaPV in three individual plants (19P-148, 19P-149 and 19P-150), while MMaV was found to originate from a single source (19P-176). Most of the virus-infected samples detected in this study were sourced from smallholdings and volunteer sites with only two of the twelve virus-positive pools having been sampled from commercial maize fields.

2.3.2 RNA sequencing and *de novo* assembly

To confirm the presence of MaPV and MMaV, RNA extracts from individual plant samples 19P-149 (positive for MaPV) and 19P-176 (MMaV-positive) were subjected to NGS. After quality trimming the NGS data, sample 19P-149 yielded 1,302,261 reads and sample 19P-176 yielded 1,196,971 reads, both with an average length of 118 bp and were subjected to *de novo* assembly. Sample 19P-149 produced a total of 548 contigs with a size range of 479-9,022 nt, from which 13 had high identity to known viruses when analysed with BLASTn, while only two virus-like contigs were identified for sample 19P-176 from a total of 753 contigs with a size range of 463 to 11,992 nt.

Contigs with high identity (≥ 99%) to MMaV and MaPV variants were identified as expected (Table 2.3). Additional contigs with high identity (91.6-98.9%) to other viruses were also detected, including two distinct MATV variants, maize-associated partiti-like virus (available on GenBank but not yet discussed in literature), maize stripe virus (MStV) and two Zea mays chrysovirus 1 (ZMCV1) strains, namely 63 (ZMCV1-63) and 201 (ZMCV1-201) (Table 2.3). No MCMV-, potyvirid- or MSV-like contigs were identified from the reads produced for either of the individual plant samples.

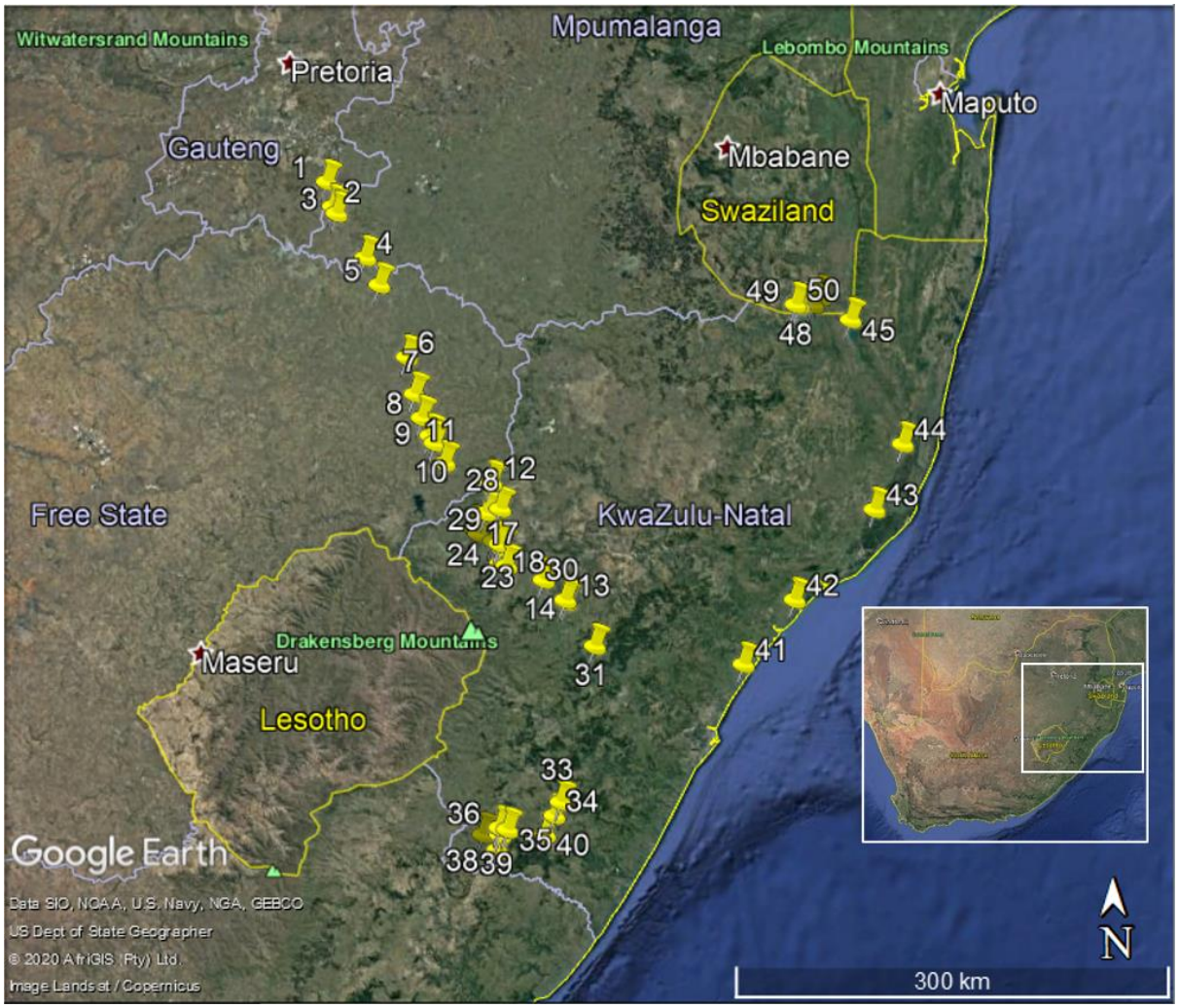


Fig. 2.1 Sites with virus-like symptoms sampled along the major maize grain transport route in KwaZulu-Natal, South Africa. (Google Earth Pro 7.3.3.7786 2015).

Table 2.2 Results of pooled maize samples tested for five different maize-infecting viruses using PCR/RT-PCR. If a virus was detected it is indicated by "✓", while "-" indicates that the virus was not detected in the sample pool.

Sample pool	Virus status				
	MSV	MaPV	MMaV	MATV-1-Tanz	MATV-4-Tanz
19P-1 to -5	-	-	-	-	-
19P-6 to -10	-	-	-	-	-
19P-11 to -15	-	-	-	-	-
19P-16 to -20	-	-	-	-	-
19P-21 to -25	-	-	-	-	-
19P-26 to -30	-	-	-	-	-
19P-31 to -35	-	-	-	-	-
19P-36 to -40	-	-	-	-	-
19P-41 to -45	-	-	-	-	-
19P-46 to -50	-	-	-	-	-
19P-51 to -55	-	-	-	-	-
19P-56 to -60	✓	-	-	-	-

19P-61 to -65	-	-	-	-	-
19P-66 to -70	✓	-	-	-	-
19P-71 to -75	✓	-	-	-	-
19P-76 to -80	✓	-	-	-	-
19P-81 to -85	-	-	-	-	-
19P-86 to -90	-	-	-	-	-
19P-91 to -95	-	-	-	-	-
19P-96 to -98	-	-	-	-	-
19P-101 to -105	-	-	-	-	-
19P-106 to -110	-	-	-	-	-
19P-111 to -115	-	-	-	-	-
19P-116 to -120	-	-	-	-	-
19P-121 to -125	-	-	-	-	-
19P-126 to -130	-	-	-	-	-
19P-131 to -135	-	-	-	-	-
19P-136 to -140	-	-	-	-	-
19P-141 to -145	-	-	-	-	-
19P-146 to -150	✓	✓	-	-	-
19P-151 to -155	-	-	-	-	-
19P-156 to -160	✓	-	-	-	-
19P-161 to -165	-	-	-	-	-
19P-166 to -170	✓	-	-	-	-
19P-171 to -175	✓	-	-	-	-
19P-176 to -180	✓	-	✓	-	-
19P-181 to -185	-	-	-	-	-
19P-186 to -190	✓	-	-	-	-
19P-191 to -195	✓	-	-	-	-
19P-196 to -200	-	-	-	-	-
19P-201 to -205	-	-	-	-	-
19P-206 to -210	✓	-	-	-	-

Table 2.3 Query coverage and nucleotide identity of *de novo*-assembled contigs to BLASTn references with the best overall scores (all E-values = 0.0). *De novo* assembly performed on CLC Genomics Workbench 20.0.2 (https://digitalinsights.qiagen.com) with length and similarity fractions of 0.9.

Sample	BLASTn hit					
	Description	Isolate	Accession	Length (nt)	Query coverage (%)	Nucleotide identity (%)
19P-149	MaPV RNA1	SSF4	MF372912	5741	100.0	99.9
	MaPV RNA2	16_0060	MK112503	2713	99.5	99.9
	MATV3	T2F2S4	MF425849	4987	100.0	96.7
	MATV-4-Tanz	16-0130b	MK066243	5583	99.6	97.8
	MStV RNA1	-	AJ969411	1413	97.2	98.3
	MStV RNA2	-	NC_038751	3337	100.1	94.1
	MStV RNA3	Tirupati	JN579656	2357	56.2	88.1
	MStV RNA4	-	NC_038752	2227	68.7	89.8
	MStV RNA5	Kurnool	JN626912	1317	99.9	96.5
	Maize-associated partiti-like virus	SS9A	MF372918	1849	100.0	98.9

	ZMCV1-63	-	MH931203	2718	26.1	94.9
	ZMCV1-201	-	MH931198	3624	16.1	91.6
19P-176	MMaV	16-0112	MK063878	12185	98.4	99.0

2.3.3 Reference mapping

Trimmed reads were then mapped to relevant reference sequences for each of the identified viruses. The mapped reads covered 68.6 to 100.0% of the lengths of the reference sequences and the average coverage depth obtained varied from as little as 3 X for the ZMCV1 variants, to as much as 10,301 X for the MMaV variant (Table 2.4). Trimmed reads were also mapped to reference sequences for MCMV (NC_003627), SCMV (NC_003398), MDMV (NC_003377), JGMV (NC_003606), WSMV (NC_001886) and MSV (HQ693424). However, few to no reads successfully mapped to these reference sequences, with reads covering less than 10% of the respective reference lengths and having an average coverage depth of less than 1 X.

Table 2.4 The percentage of reference length mapped, and the average depth of coverage obtained for a variety of reference-mapped maize virus sequences. References with less than 1 X average coverage depth were omitted.

Sample	Mapping reference				Reference-mapped consensus seqs	
	Description	Isolate	Accession	Length (nt)	Reference length coverage (%)	Average coverage depth (X)
19P-149	MaPV RNA1	16_0060	MK112502	5806	100.0	286
	MaPV RNA2	16_0060	MK112503	2713	100.0	2119
	MATV3	T2F2S4	MF425849	4987	100.0	477
	MATV-4-Tanz	16-0130b	MK066243	5583	99.6	35
	MStV RNA1	-	AJ969411	1413	100.0	2221
	MStV RNA2	-	NC_038751	3337	100.0	1423
	MStV RNA3	Tirupati	JN579656	2357	79.5	1259
	MStV RNA4	-	NC_038752	2227	91.0	2425
	MStV RNA5	Kurnool	JN626912	1317	99.9	3205
	Maize-associated partiti-like virus	SS9A	MF372918	1849	100.0	130
	ZMCV1-63	-	MH931203	2718	77.1	3
	ZMCV1-201	-	MH931198	3624	68.6	3
19P-176	MMaV	16-0112	MK063878	12185	100.0	10301

2.3.4 Consensus sequence alignments

The *de novo*-assembled contigs and the reference-mapped consensus sequences were then aligned, and partial/complete draft genome sequences extracted for BLASTn analysis (Table 2.5). A complete MMaV draft genome with 99.0% nucleotide identity from sample 19P-176, and complete/near complete MaPV RNA1 and RNA2 draft genome segments with 99.8 and 99.9% nucleotide identity to Tanzanian isolates from sample 19P-149, were assembled and submitted to GenBank with the following accession numbers: MMaV: MW063124; MaPV RNA1: MW063117; and MaPV RNA2: MW063118.

Two complete/near complete draft CDSs were produced for distinct MATV variants present in sample 19P-149, with 96.8 and 98.6% nucleotide identity to known isolates from Rwanda and were submitted to GenBank with the following accession numbers: MW063115 and MW063116. These two variants shared only 41.3% pairwise nucleotide identity. Partial draft CDSs were assembled for the variants most closely related to ZMCV1-63 and ZMCV1-201, having 93.4 and 91.3% nucleotide identity to known isolates from the USA. These sequences were submitted to GenBank with the accession numbers MW063135 and MW063136. A partial CDS with 98.9% nucleotide identity to a maize-associated partiti-like virus was also deposited into the GenBank database with accession number MW063112.

A complete/near complete set of draft sequences for the four recognised RNA segments of MStV (RNA 2-4) were assembled and shared 87.7 to 96.5% identity to other known isolates. In addition to these, a 9022 nt sequence that shared 74.1% nucleotide identity to the complete RNA1 sequence of rice stripe virus (RSV) isolate CY-CN RNA1 (GQ229089) and 98.3% nucleotide identity to the only known MStV RNA 1 partial CDS (1413 nt; AJ969411) (not shown in table) was also produced. The presence of this virus was confirmed using the following detection primers that were designed during this study: MStV-RNA4.05 F: 5′- TGG ATC AAC AGC TAT GCA GAA -3′; and MStV-RNA4.05 R: 5′- AAG TCA GGG CAT ATC ATT GTG A -3′. The 333 bp amplicon produced was Sanger sequenced bidirectionally, and its identity confirmed as MStV using BLASTn analysis.

Table 2.5 BLASTn hits with best bit-scores obtained for the *de-novo*/reference-mapped consensus sequences produced during this study. All hits had an E-value of 0.0 and query coverage ≥99% unless specified: [A] = 80%; [B] = 87%; [C] = 83%; and [D] = 85%. Draft sequence lengths represent the number of defined nucleotides per sequence.

Sample	BLASTn hit					Draft sequence	
	Description	Isolate	GenBank accession	Length (nt)	Identity (%)	GenBank accession	Length (nt)
19P-149	MaPV RNA1	16_0060	MK112502	5806	99.8	MW063117	5807
	MaPV RNA2	16_0060	MK112503	2713	99.9	MW063118	2698
	MATV	Rwanda 6	MN428833	5030	96.8	MW063115	4988
	MATV	Rwanda 3	MN428830	5579	98.6	MW063116	5563
	RSV RNA1	CY-CN	GQ229089	8970	74.1[A]	MW063119	9022
	MStV RNA2	-	NC_038751	3337	94.1	MW063120	3336
	MStV RNA3	Tirupati	JN579656	2357	87.7[B]	MW063121	2067
	MStV RNA4	-	NC_038752	2227	89.0	MW063122	2187
	MStV RNA5	Kurnool	JN626912	1317	96.5	MW063123	1316
	Maize-associated partiti-like virus	SS9A	MF372918	1849	98.9	MW063112	1849
	ZMCV1-63	-	MH931203	2718	93.4[C]	MW063135	2242
	ZMCV1-201	-	MH931198	3624	91.3[D]	MW063136	2571
19P-176	MMaV	16-0112	MK063878	12185	99.0	MW063124	12185

2.3.5 Additional MMaV confirmation

Bands of expected size were produced for each of the two additional MMaV primer sets and BLASTn analysis of the bidirectional Sanger sequencing results confirmed the identity of each as MMaV.

2.3.6 Phylogenetic analysis

The MEGA X protein model test determined that the Le Gascuel (LG) 2008 model with both Gamma distribution (n=5) and invariant sites best suited both the MaPV and MMaV data sets but that frequencies (+F) should also be taken into account for the MMaV dataset. In terms of the nucleotide model test, the Hasegawa-Kishino-Yano (HKY) model best suited the MaPV dataset, while the General Time Reversible (GTR) model best suited both the MMaV and MATV datasets, with evolutionarily invariable sites (+I) allowed for MMaV, and a discrete Gamma distribution (+G; 5 categories) employed for MATV.

Based on the amino acid sequences (Fig. 2.2A), the MaPV variant detected during this study (isolate 19P-149) grouped with other previously known MaPV variants with the tree topology being the same as that reported previously in the literature (Read et al. 2019b), with the addition of an isolate since reported from Rwanda

(GenBank accession: MN248736). However, the branching pattern observed based on the nucleotide sequences (Fig. 2.2B) suggested that the known variants may cluster somewhat separately from the MaPV variant detected in this study.

Based on the amino acid sequences (Fig. 2.3A), the MMaV variant (isolate 19P-176) also grouped with other known MMaV variants with the tree topology also being very similar to that reported previously (Read et al. 2019c). Furthermore, based on the nucleotide sequences (Fig. 2.3B), the MMaV variant clustered most closely with Tanzanian isolate 16-0112.

The tree topology produced for the MATV nucleotide sequences (Fig. 2.4) was comparable to that described in previous literature (Read et al. 2019a). Isolate 19P-149a clustered closely with isolates from Kenya, Rwanda, Ethiopia, South Sudan, Tanzania and China mainly classified as MATV3, while isolate 19P-149b clustered with more recently reported isolates from China, Rwanda and Tanzania and classified as MATV but clustered distantly from other previously classified MATV types (MATV1, MATV2 and MATV3) (Fig. 2.4).

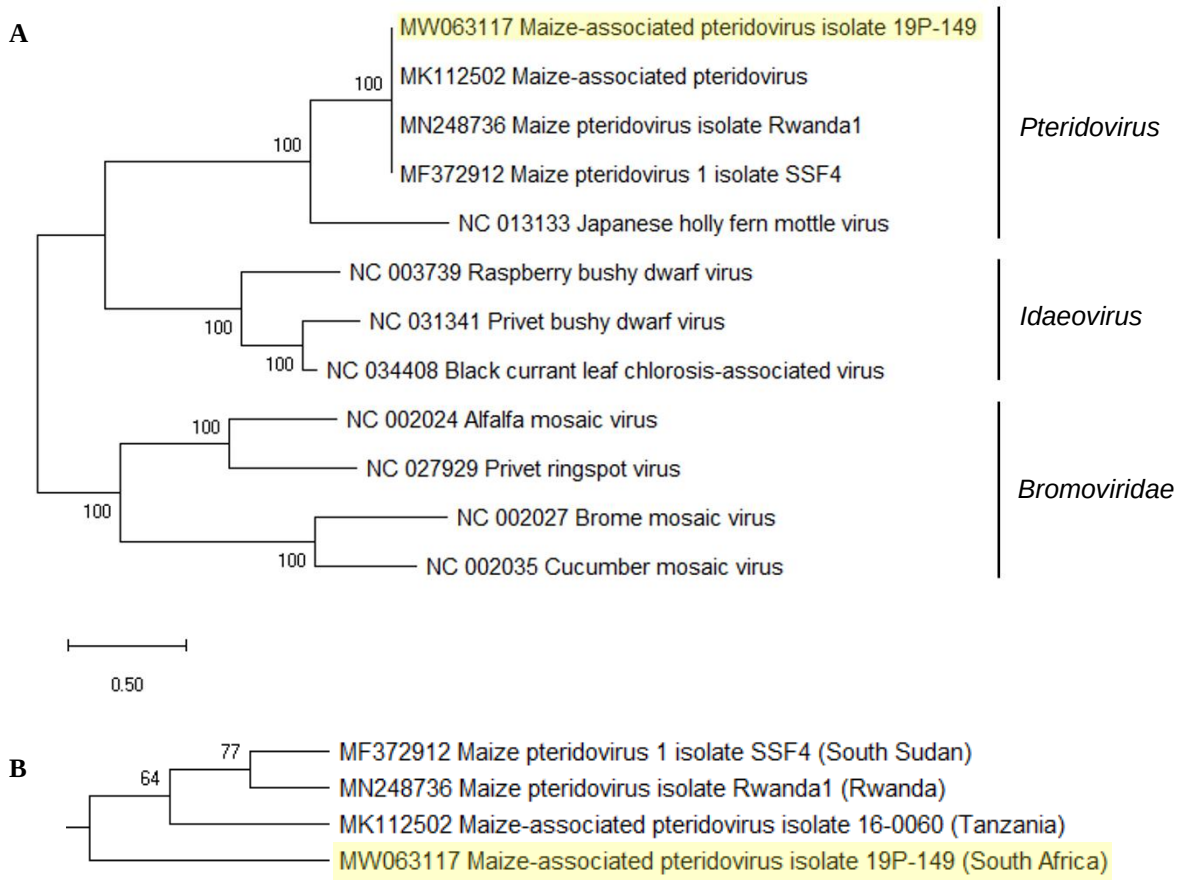


Fig 2.2 Maximum Likelihood trees showing the relationship of the maize-associated pteridovirus (MaPV) variant produced during this study (highlighted in yellow) against other known variants and constructed with 1,000 bootstrap replicates. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test shown next to the branches. **(A)** Phylogram based on the amino acid (aa) sequences of the RNA-dependent RNA polymerase (RdRp) domain of MaPV variants and the cognate region of eight other related viruses available on GenBank. The phylogram was constructed in MEGA X (Kumar et al. 2018) using the Le Gascuel 2008 model with discrete Gamma distribution (+G; 5 categories), allowing for evolutionarily invariable sites (+I). The bar indicates the numbers of substitutions per site. **(B)** Cladogram based on the complete RNA1 genome sequences of all known MAPV variants. The cladogram was constructed using the Hasegawa-Kishino-Yano (HKY) model and rooted using the Japanese holly fern mottle virus (GenBank accession: NC013133).

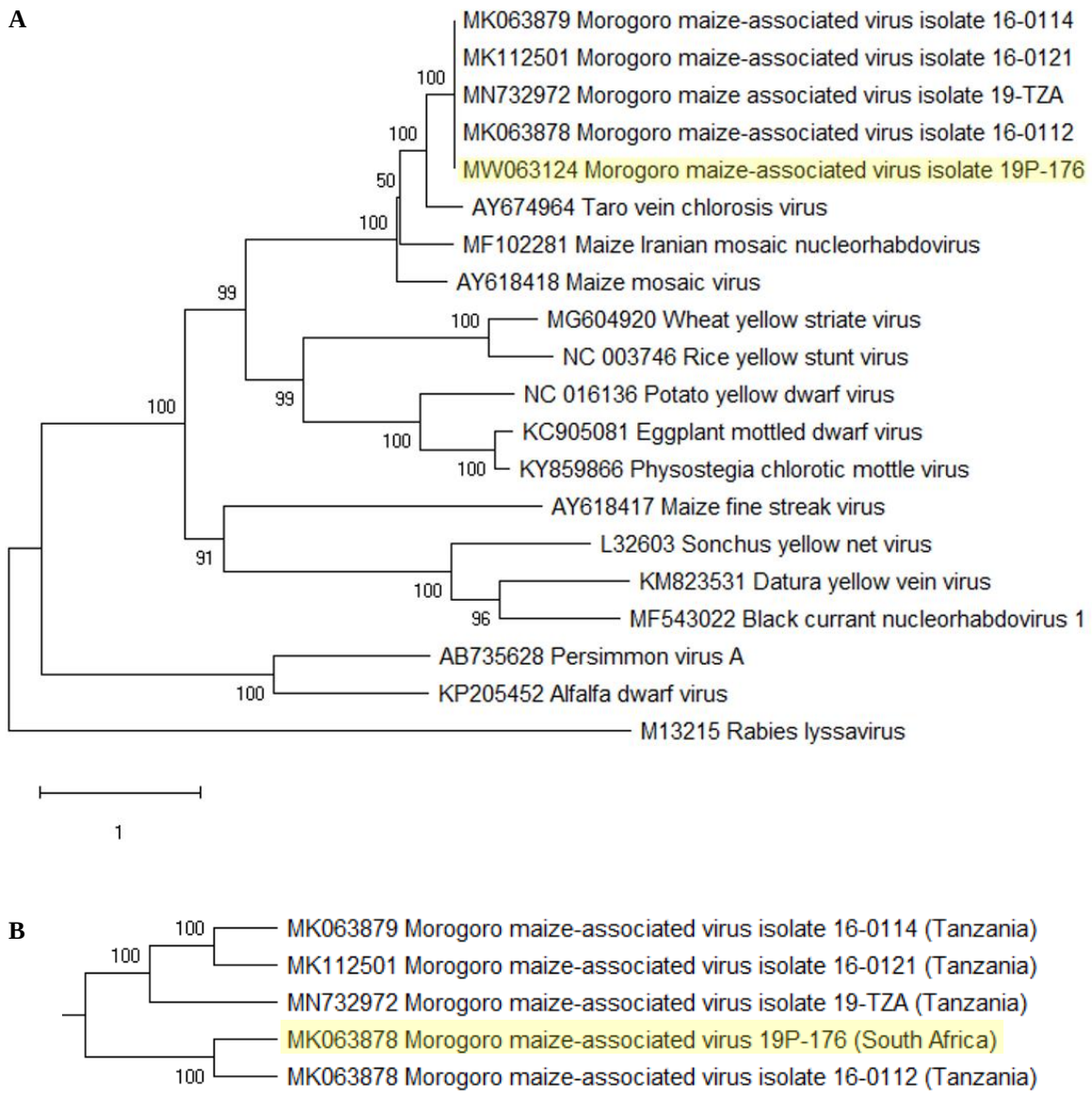


Fig 2. 3 Maximum Likelihood trees showing the relationship of the Morogoro maize-associated virus (MMaV) variant produced during this study (highlighted in yellow) against other known variants and constructed with 1,000 bootstrap replicates. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test shown next to the branches. **(A)** Phylogram based on the amino acid (aa) sequences of the cognate L-protein domain of several other plant-infecting members of the family *Rhabdoviridae*. The phylogram was constructed in MEGA X (Kumar et al. 2018) using the Le Gascuel 2008 model with frequencies (+F), discrete Gamma distribution (+G; 5 categories) and allowing for evolutionarily invariable sites (+I). The bar indicates the numbers of substitutions per site. **(B)** Cladogram based on the complete genome sequences of all known MMaV variants. The cladogram was constructed using the General Time Reversible (GTR) model, allowing for evolutionarily invariable sites (+I) and rooted using the maize Iranian mosaic nucleorhabdovirus (GenBank accession: MF102281).

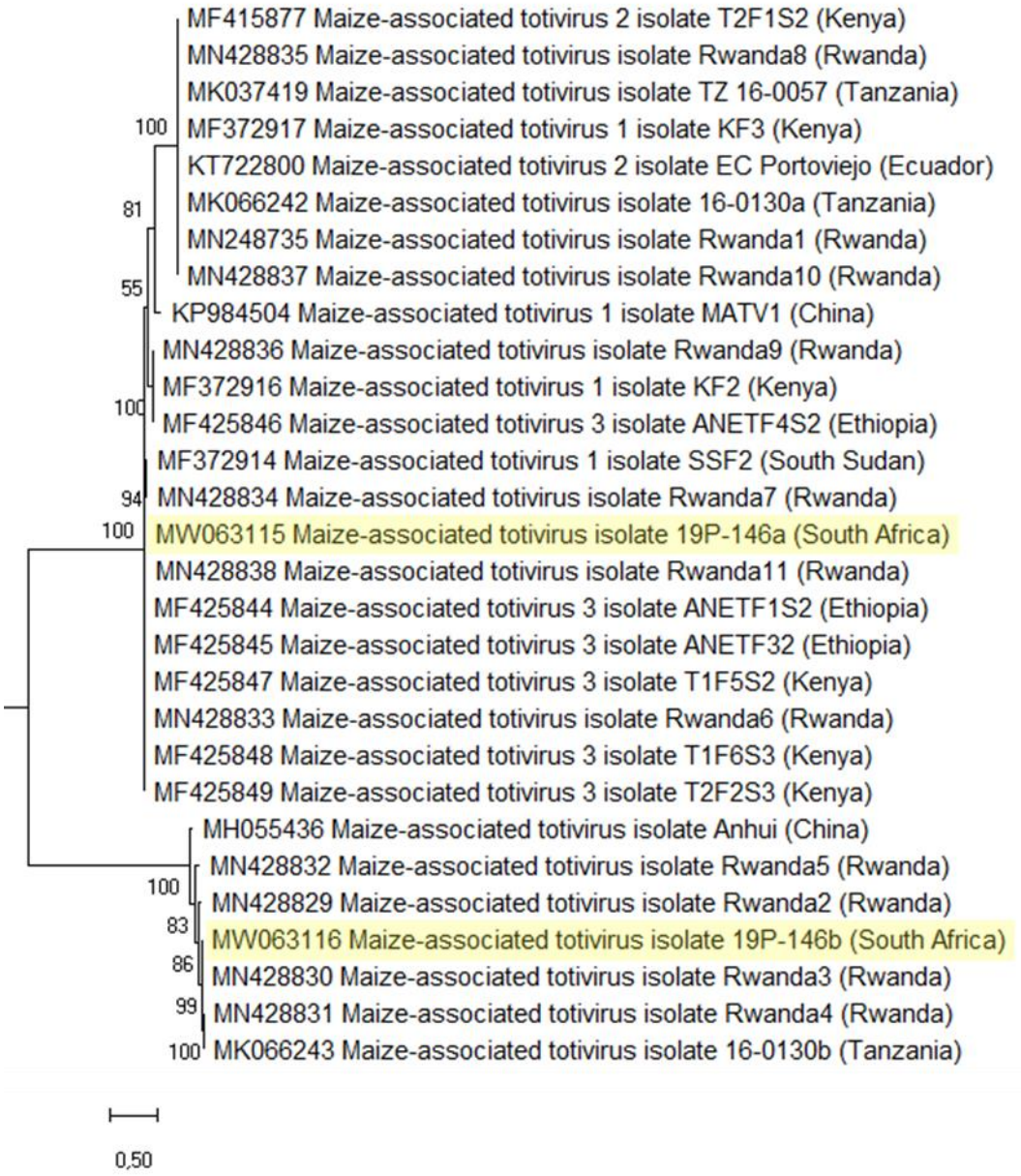


Fig 2.4 Maximum Likelihood phylogram based on the complete coding domain sequence of the RNA-dependent RNA polymerase (RdRp) gene of all known maize-associated totivirus (MATV) variants, including the variants produced during this study. The phylogram was constructed in MEGA X (Kumar et al. 2018) using the General Time Reversible (GTR) model with discrete Gamma distribution (+G; 5 categories) and 1,000 bootstrap replicates. The tree was rooted using the black raspberry virus F (GenBank accession: EU082131). The percentage of trees in which the associated taxa clustered together is shown next to the branches. A total of 5,557 nucleotide positions across the 30 sequences were included in the final dataset. The MATV sequences produced during this study are highlighted in yellow. The bar indicates the numbers of substitutions per site.

2.4 Discussion

Viruses were detected in 13 sample pools (10 sites) amongst the 42 sample pools (39 sites) tested, with infected samples originating primarily from smallholdings and volunteer sites. No instances of MCMV or potyvirids,

which may indicate predisposition to MLND, were detected during this survey. This was unexpected, as potyviruses SCMV and JGMV have been reported in South African maize previously (Schulze 2018). However, only a few areas sampled during this study overlapped with regions where potyviruses were previously detected (Schulze 2018). Furthermore, during the 2018-2019 season, South Africa faced extreme dry conditions. This led to farmers sowing their fields later than usual, likely leaving the land barren for longer than usual, potentially reducing the availability of alternate hosts and presenting an unfavourable environment for vector survival, thus potentially affecting the persistence of potyviruses between seasons.

MSV was found to be the most common and widespread virus affecting maize in KwaZulu-Natal, supporting previous reports (Sibiya et al. 2013). Two viruses recently discovered in maize, MaPV and MMaV, were detected by RT-PCR and confirmed with NGS. Complete/near complete MaPV draft genome segment sequences and a complete MMaV draft genome sequence were constructed using both *de-novo* and reference based assembly and submitted to GenBank (accessions: MaPV RNA1 isolate 19P-149: MW063117; MaPV RNA2 isolate 19P-149: MW064118, and MMaV isolate 19P-179: MW063124). The nine terminal nucleotides on either end of the genome displayed inverted complementarity as previously reported for MMaV and other rhabdoviruses (Jackson et al.2005; Read et al. 2019a). Based on the RdRp regions of these viruses, both MaPV and MMaV variants clustered closely to other known virus variants (Fig. 2.2A and Fig. 2.3A, respectively). The MaPV variant (isolate 19P-149) may be a somewhat diverged variant as it branched away from those Tanzania, South Sudan, and Rwanda (Fig. 2.2B). The branching pattern observed for the MMaV variant (19P-176), on the other hand, suggests it may be most closely related to Tanzanian isolate 16-0112 (Fig. 2.3B). This is the first time either of these viruses have been reported in South Africa. Since MMaV had not been found outside of Tanzania previously, the presence of this virus was further confirmed using two additional primer sets.

The highest prevalence of viruses was detected from smallholder farms near the Eswatini (formerly Swaziland) border (Table 2.2 and Supplementary 2.1). This is of concern as one of the predicted routes for the spread of MCMV to South Africa is through this north eastern part of South Africa. There are a variety of factors that may be responsible for plants from smallholder farms being more vulnerable to virus infections than those from commercial farms. Since plants from smallholders are primarily sourced from seed retained from the previous season, and continuous maize planting may occur without the implementation of rotational cropping strategies, viruses introduced during a growing season are likely to persist through to the following growing season. Furthermore, it is likely that commercial farms experience lower levels of stress than smallholder farms due to better irrigation, fertiliser/rotational cropping practices and insecticide use. Additionally, transgenic Bt-gene containing maize grown by commercial farmers may also be influencing the prevalence of insects other than stem borers for which the strain was initially developed.

During this study, complete/near complete draft CDSs were assembled for two distinct MATV variants with nucleotide identities between the two variants being relatively low. One of the MATV variants (19P-149a; GenBank accession: MW063115) was most similar to isolate Rwanda6, believed to be a member of the group MATV3, while the other MATV variant (19P-149b; GenBank accession: MW063116) was most similar to isolate Rwanda3 and the divergent MATV-4-Tanz isolate from Tanzania (Table 2.5). These relationships were supported by the clustering pattern observed in the MATV Maximum Likelihood tree (Fig. 2.4). Due to isolate 19P-149b clustering with some new MATV isolates of unclassified groups that clustered far away from MATV1, MATV2 and MATV3 isolates, and with isolate 19P-149b sharing as little as ~41% pairwise nucleotide identity with isolate

19P-149a, it is likely that isolate 19P-149b is a representative of a new group of MATV variants, tentatively named MATV4.

Due to the taxonomy of MATV variants remaining unresolved at present, it is currently unclear whether the different MATV groups represent different orders or genera of the family *Totiviridae*. It is possible that some MATV sequences available on GenBank may have been misclassified. The phylogenetic tree suggested that some of the MATV1-classified sequences clustered more closely with MATV2 or MATV3 representatives (Fig. 2.4). These sequences may need to be reclassified in future as more sequence data becomes available and the phylogenies better resolved.

The RT-PCR based survey (Table 2.2) failed to detect the presence of the two MATV variants detected using NGS in sample 19P-149 probably due to the presence of nucleotide variation observed by NGS in the PCR primer-binding regions. This confirms the potential of NGS in novel virus detection while highlighting the limitations of PCR-based diagnostics. To improve future detection of a greater range of MATV variants using RT-PCR, primers should be designed over more conserved regions of the genome. If no suitable conserved regions are identified among all known MATV variants of the family *Totiviridae*, it is suggested that genus/species-specific primer sets may be designed based on the different MATV groups identified thus far.

It is currently unclear whether MATV infects plants or whether it infects fungi living in/on the plant. Although, since most other members of the family *Totiviridae* are mycoviruses, it is more likely that MATV would also be a mycovirus (Asiimwe et al. 2020; Chen, Cao, et al. 2016). Member of the family *Chrysoviridae* (including ZMCV1) are also believed to be mycoviruses rather than plant-infecting viruses, although plant and possibly insect-infecting mycoviruses have been reported (Kotta-Loizou et al. 2020). ZCMV1 was recently identified for the first time in ancient maize cobs from the USA, and has since also been reported in maize leaf samples from Rwanda (Asiimwe et al. 2020). During the present study, bioinformatic analysis of the NGS data also resulted in the assembly of partial draft CDSs for two ZCMV1 strain variants, namely ZMVC1-63 and ZMCV1-201 (GenBank accessions: MW063135 and MW063136, respectively). The average coverage depth produced for these sequences was very low compared to the other viruses detected (Table 2.4), indicating that there may have been low virus titres in the plant.

MStV is a tenuivirus transmitted in a persistent manner by the planthopper *Peregrinus maidis* (family *Delphacidae*) (Nault & Ammar 1989) and has been reported in many tropical and sub-tropical areas in Africa (Thottappilly et al. 1993), the Americas (Nault 1979; Tsai 1975), and Australia (Greber 1981). The MStV vector has been reported in South Africa (Muir 1929; Thottappilly et al. 1993), and MStV is believed to be present (EPPO Global Database 2015), although this is not based on any published study to our knowledge. Symptoms of MStV include extreme stunting, broad yellow-light green striping or complete yellowing of leaves, apical bending, necrosis, and dieback (Kulkarni 1973; Thottappilly et al. 1993). This genome of this virus has been reported to consist of five RNA segments (RNA1-5) and, although MStV was first reported in 1936 (Storey 1936), prior to the present study, no complete RNA1 sequences and only two to four complete RNA2-5 sequences were available on GenBank.

In this study, complete/near complete sequences for MStV segments RNA2-5 (GenBank accessions: MW063120-MW063123) were produced, along with an RNA1 sequence (GenBank accession: MW063119) having ~74% nucleotide identity to RSV and ~98% identity to a MStV RNA1 partial sequence (Table 2.5). Thus, the MStV RNA1 sequence assembled during this study is likely the most complete sequence constructed for the

MStV RNA1 segment to date. The presence of this virus was confirmed via RT-PCR using primers developed during this study. This is the first time MStV has been confirmed in South Africa, with the infected plant exhibiting yellowing over a large portion of the leaf (Supplementary 2.1), a known symptom of MStV (Kulkarni 1973). Since some of the segments of the MStV variant had relatively low nucleotide identity to known isolates (as low as ~74%), overlapping primers have been designed across the entire MStV genome which will be used to confirm the segment sequences as part of future research. Once this has been achieved, the presence of MStV for the first time in South Africa, along with its genome sequence, will be submitted for publication and the draft genome sequences on GenBank updated if necessary.

Considering MStV was first identified around 84 years ago (Storey 1936), reports published on this virus to date appear to be relatively few, with most MStV-related works published between the 1970s and 1990s. One study published in 1992, reported mixed infections of MSV and MStV in a susceptible maize variety, ZS 5206, in Mauritius (Roca de Doyle and Autrey 1992). The report suggested that single infections of MStV produced yield losses greater than those observed for single infections of MSV, with yield losses averaged across 10 biological replicates for each type of infection tested at two separate growth stages (at 3-5 leaf stage and at 7-10 leaf stage) (Roca de Doyle and Autrey 1992). Furthermore, co-infections of MStV and MSV produced yield losses greater than single infections of either virus. This may suggest that co-infection of the viruses may additively affect the maize plant, thus producing more severe symptoms than individual infections. However, additional studies potentially using greater numbers of replicates should be used to determine whether similar results are obtained from other maize varieties. Roca de Doyle and Autrey (1992) also reported that symptoms of MStV infections were masked by MSV-like symptoms when the viruses co-occurred in maize. This means that yield losses may have been incorrectly attributed solely to MSV in the past if adequate MStV detection methods were not employed, and the incidence of MStV potentially underestimated (Roca de Doyle and Autrey 1992). Thus, it is suggested that detection methods of both MSV and MStV be used in parallel in future when symptoms synonymous of MSV infection are observed.

In conclusion, MSV continues to be the most important virus affecting maize in South Africa, with infections detected in 29% of sample pools tested. The presence of MaPV and MMaV were reported and MStV confirmed for the first time in South Africa during this study with the detection of MATV, ZMCV1-63 and ZMCV1-201 as preliminary findings. Virus infections were detected primarily from smallholder farmers with the highest virus prevalence occurring in northern KwaZulu-Natal, near the Eswatini border, raising concerns around the possible introduction in future of MCMV to South Africa through this northern KwaZulu-Natal region which also borders Mozambique. Further research is required to determine the effects MaPV and MMaV may have on maize and maize yields, their possible vectors, and whether they are seed-transmissible in maize. In order to determine the risk, if any, these viruses may pose to South Africa's maize industry, the incidence and distribution of MMaV, MaPV, and MStV (as well as those reported as preliminary findings) across other maize-growing regions of the country should be investigated. Although no instances of MCMV were detected during this study, it is advised that surveillance for MCMV in South Africa continue so that early detection may enable the implementation of efficient disease management procedures to prevent a possible outbreak.

2.5 References

Adams, I. P., Harju, V. A., Hodges, T., Hany, U., Skelton, A., Rai, S., et al. (2014). First report of maize lethal necrosis disease in Rwanda. *New Disease Reports*, 29, 22.

Allsopp, E. (2010). Investigation into the apparent failure of chemical control for management of western flower thrips, Frankliniella occidentalis (Pergande), on plums in the Western Cape Province of South Africa. *Crop Protect*ion, 29(8), 824–831.

Asiimwe, T., Stewart, L. R., Willie, K., Massawe, D. P., Kamatenesi, J., & Redinbaugh, M. G. (2020). Maize lethal necrosis viruses and other maize viruses in Rwanda. *Plant Pathology*, 69(3), 585–597.

Brault, V., Uzest, M., Monsion, B., Jacquot, E., & Blanc, S. (2010). Aphids as transport devices for plant viruses. *Comptes Rendus Biologies*, 333(6–7), 524–538.

Cabanas, D., Watanabe, S., Higashi, C. H. V, & Bressan, A. (2013). Dissecting the mode of maize chlorotic mottle virus transmission (Tombusviridae: Machlomovirus) by Frankliniella williamsi (Thysanoptera: Thripidae). *Journal of Economic Entomology*, 106(1), 16–24.

Chen, S., Cao, L., Huang, Q., Qian, Y., & Zhou, X. (2016). The complete genome sequence of a novel maize-associated totivirus. *Archives of Virology*, 161(2), 487–490.

EPPO Global Database. (2015). Maize stripe tenuivirus. https://gd.eppo.int/taxon/MSPV00/distribution. Accessed 18 November 2020

FAO REOA. (2013). Maize lethal necrosis disease (MLND) - A snapshot. http://www.fao.org/fileadmin/user_upload/emergencies/docs/MLND Snapshot_FINAL.pdf. Accessed 25 November 2020

Flett, B., & Mashingaidze, K. (2016). Maize lethal necrosis: Possible threat to local maize production. *Grain SA*. http://www.grainsa.co.za/maize-lethal-necrosis:-possible-threat-to-local-maize-production. Accessed 13 March 2020

Google Earth Pro 7.3.3.7786. (2015). KwaZulu-Natal, South Africa, 28° 27’ 33.48’’S, 30° 10’ 27.99’’E, eye alt 673.37 km. https://www.google.com/earth/index.html. Accessed 19 October 2020

Greber, R. S. (1981). Maize stripe disease in Australia. *Australian Journal of Agricultural Research*, 32(1), 27–36.

Hatting, J. L., Humber, R. A., Poprawski, T. J., & Miller, R. M. (1999). A survey of fungal pathogens of aphids from South Africa, with special reference to cereal aphids. *Biological Control*, 16(1), 1–12.

Isabirye, B. E., & Rwomushana, I. (2016). Current and future potential distribution of maize chlorotic mottle virus and risk of maize lethal necrosis disease in Africa. *Journal of Crop Protection*, 5(2), 215–228.

Jackson, A. O., Dietzgen, R. G., Goodin, M. M., Bragg, J. N., & Deng, M. (2005). Biology of plant rhabdoviruses. *Annual Review of Phytopathology*, 43, 623–660.

Jensen, S. G. (1985). laboratory transmission of maize chlorotic mottle virus by three species of corn rootworms. *Plant Disease*, 69(10), 864–868.

Jensen, S. G., Wysong, D. S., Ball, E. M., & Higley, P. M. (1991). Seed transmission of maize chlorotic mottle virus. *Plant Disease*, 75(5), 497–498.

Kagoda, F., Gidoi, R., & Isabirye, B. E. (2016). Status of maize lethal necrosis in eastern Uganda. *African Journal of Agricultural Research*, 11(8), 652–660.

King, A. M. Q., Lefkowitz, E., Adams, M. J., & Carstens, E. B. (2011). *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses (Vol. 9).* San Diego: Elsevier.

Kiruwa, F. H., Feyissa, T., & Ndakidemi, P. A. (2016). Insights of maize lethal necrotic disease: A major constraint to maize production in East Africa. *African Journal of Microbiology Research*, 10(9), 271–279.

Kotta-Loizou, I., Castón, J. R., Coutts, R. H. A., Hillman, B. I., Jiang, D., Kim, D.-H., et al. (2020). ICTV virus taxonomy profile: Chrysoviridae. *Journal of General Virology*, 101(2), 143.

Kriticos, D. J., Reynaud, P., Baker, R. H. A., & Eyre, D. (2012). Estimating the global area of potential establishment for the western corn rootworm (Diabrotica virgifera virgifera) under rain-fed and irrigated agriculture. *EPPO Bulletin*, 42(1), 56–64.

Kulkarni, H. Y. (1973). Comparison and characterization of maize stripe and maize line viruses. *Annals of Applied Biology*, 75(2), 205–216.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549.

Li, L., Wang, X., & Zhou, G. (2011). Effects of seed quality on the proportion of seed transmission for Sugarcane mosaic virus in maize. *Cereal Research Communications*, 39(2), 257–266.

Lukanda, M., Owati, A., Ogunsanya, P., Valimunzigha, K., Katsongo, K., Ndemere, H., & Kumar, P. L. (2014). First report of maize chlorotic mottle virus infecting maize in the Democratic Republic of the Congo. *Plant Disease*, 98(10), 1448.

Mahuku, G., Lockhart, B. E., Wanjala, B., Jones, M. W., Kimunye, J. N., Stewart, L. R., et al. (2015). Maize lethal necrosis (MLN), an emerging threat to maize-based food security in sub-Saharan Africa. *Phytopathology*. https://doi.org/10.1094/PHYTO-12-14-0367-FI

Mahuku, G., Wangai, A., Sadessa, K., Teklewold, A., Wegary, D., Ayalneh, D., et al. (2015). First report of maize chlorotic mottle virus and maize lethal necrosis on maize in Ethiopia. *Plant Disease*, 99(12), 1870–1870.

Muir, F. (1929). XVI.—New and little-known African Delphacidæ (Homoptera, Fulgoroidea) in the collection of the British Museum. *Annals and Magazine of Natural History*, 4(20), 186–222.

Nault, L. R. (1979). Identification of Maize Viruses and Mollicutes and Their Potential Insect Vectors in Peru. *Phytopathology*, 69(8), 824. https://doi.org/10.1094/phyto-69-824

Nault, L. R., & Ammar, E. D. (1989). Leafhopper and planthopper transmission of plant viruses. *Annual Review of Entomology*, 34(1), 503–529.

Nault, L. R., Styer, W. E., Coffey, M. E., Gordon, D. T., Negi, L. S., & Niblett, C. L. (1978). Transmission of maize chlorotic mottle virus by chrysomelid beetles. *Phytopathology*, 68, 1071–1074.

Pratt, C. F., Constantine, K. L., & Murphy, S. T. (2017). Economic impacts of invasive alien species on African smallholder livelihoods. *Global Food Security*, 14(November 2016), 31–37.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019a). Characterization and detection of maize-associated pteridovirus (MaPV), infecting maize (Zea mays) in the Arusha region of Tanzania. *European Journal of Plant Pathology*, 154, 1165–1170.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019b). Molecular characterization of Morogoro maize-associated virus, a nucleorhabdovirus detected in maize (Zea mays) in Tanzania. *Archives of Virology*, 164, 1711–1715.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019c). Diversity and distribution of maize-associated totivirus strains from Tanzania. *Virus Genes*, 55, 429–432.

Read, D. A., Featherstone, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019). First report of maize yellow mosaic virus (MaYMV) on maize (Zea mays) in Tanzania. *Journal of Plant Pathology*, 101(1), 203.

Redinbaugh, M. G., & Stewart, L. R. (2018). Maize lethal necrosis: An emerging, synergistic viral disease. *Annual Review of Virology*, 5(August), 301–322.

Roca de Doyle, M. M., & Autrey, L. J. C. (1992). Assessment of yield losses as a result of co-infection by maize streak virus and maize stripe virus in Mauritius. *Annals of Applied Biology*, 120(3), 443–450.

Schulze, S. E. (2018). *Pre-empting maize lethal necrosis disease in South Africa: Potyviruses of maize (Zea mays). M.Sc. Thesis*. University of Pretoria.

Shepherd, D. N., Martin, D. P., van der Walt, E., Dent, K., Varsani, A., & Rybicki, E. P. (2010). Maize streak virus: An old and complex "emerging" pathogen. *Molecular Plant Pathology*, 11(1), 1–12.

Shishkin, A. A., Giannoukos, G., Kucukural, A., Ciulla, D., Busby, M., Surka, C., et al. (2015). Simultaneous generation of many RNA-seq libraries in a single reaction. *Nature Methods*, 12, 323–325.

Sibiya, J., Tongoona, P., Derera, J., & Makanda2, I. (2013). Smallholder farmers' perceptions of maize diseases, pests, and other production constraints, their implications for maize breeding and evaluation of local maize cultivars in KwaZulu-Natal, South Africa. *African Journal of Agricultural Research*, 8(17), 1790–1798.

Somsen, H. W., & Sill, W. H. (1970). *The wheat curl mite, Aceria tulipae Keifer, in relation to epidemiology and control of wheat streak mosaic*. Manhattan: Agricultural Experiment Station.

Stewart, L. R., Teplier, R., Todd, J. C., Jones, M. W., Cassone, B. J., Wijeratne, S., et al. (2014). Viruses in maize and Johnsongrass in Southern Ohio. *Phytopathology*. https://doi.org/10.1094/PHYTO-08-13-0221-R

Storey, H. H. (1924). The transmission of a new plant virus disease by insects. *Nature*, 114(2859), 245.

Storey, H. H. (1925). The transmission of streak disease of maize by the leafhopper Balclutha mbila Naude. *Annals of Applied Biology*, 12(4), 422–439.

Storey, H. H. (1936). Virus diseases of East African plants. *The East African Agricultural Journal*, 1(4), 333–337.

Thottappilly, G., Bosquw-Perez, N. A., & Rossel, H. W. (1993). Viruses and virus diseases of maize in tropical Africa. *Plant Pathology*, 42(4), 494–509.

Tsai, J. H. (1975). Occurrence of a corn disease in Florida transmitted by Peregrinus maidis. *Plant Disease Reporte*r, 59(10), 830–833.

van den Berg, N., Crampton, B. G., Hein, I., Birch, P. R. J., & Berger, D. K. (2004). High-throughput screening of suppression subtractive hybridization cDNA libraries using DNA microarray analysis. *BioTechniques*, 37(5), 818–824.

von Wechmar, M. B., Chauhan, R., Heam, S., & Knox, E. (1987). Applications of immunoelectroblotting to differentiate between strains of maize dwarf mosaic virus and sugarcane mosaic virus occurring in South Africa. In *CSFRI Symposium:(Research into Citrus and Subtropical Crops)* Abstract 59.

Wangai, A. W., Redinbaugh, M. G., Kinyua, Z. M., Miano, D. W., Leley, P. K., Kasina, M., et al. (2012). First report of maize chlorotic mottle virus and maize lethal necrosis in Kenya. *Plant Disease*, 96(10), 1582.

Welgemoed, T., Pierneef, R., Read, D. A., Schulze, S. E., Pietersen, G., & Berger, D. K. (2020). Next generation sequencing reveals past and current widespread occurrence of maize yellow mosaic virus in South Africa. *European Journal of Plant Pathology*, 158, 237–249.

White, E. J., Venter, M., Hiten, N. F., & Burger, J. T. (2008). Modified cetyltrimethylammonium bromide method improves robustness and versatility: The benchmark for plant RNA extraction. *Biotechnology Journal*, 3(11), 1424–1428.

Willment, J. A., Martin, D. P., & Rybicki, E. P. (2001). Analysis of the diversity of African streak mastreviruses using PCR-generated RFLPs and partial sequence data. *Journal of Veterinary Medical Science*, 93, 75–87.

Zheng, L., Rodoni, B. C., Gibbs, M. J., & Gibbs, A. J. (2010). A novel pair of universal primers for the detection of potyviruses. *Plant Pathology*, 59, 211–220.

Chapter 3: Current genetic diversity and distribution of maize streak virus in South Africa and neighbouring maize-growing regions

3.1 Introduction

Since maize streak virus (MSV), the causative agent of maize streak disease (MSD), is estimated to have originated in southern Africa around 1863 (Monjane et al. 2011), crop losses of up to 100%, totalling around 280 million USD per year, have been reported in Africa (Bosque-Pérez 2000). For this reason, MSV is considered the most economically important virus affecting maize in Africa (Rybicki 2015). Prior to the present study, 11 types (MSV-A to -K) (Martin et al. 2001; Varsani et al. 2008), five MSV-A subtypes [MSV-A_1 to -A_4 and -A_6, (Martin et al. 2001) with MSV-A_5 representing a sublineage of MSV-A_1 (Owor et al. 2007)], and three MSV-B subtypes (MSV-B_1 to -B_3) (Varsani et al. 2008) had been identified.

Although MSV-resistant genotypes have been commercially available for the past 40 years (Soto et al. 1982), African countries such as Uganda have reported little progress in poverty-stricken regions with MSV persisting in smallholder farms (Owor et al. 2007). To date, only MSV-A_1 and A_4 have been reported in South Africa, with MSV-A_2 reported in West Africa, MSV-A_3 in East Africa, and MSV-A_6 endemic to the islands of La Réunion and Mauritius (Martin et al. 2001).

In 2011, a new virus disease known as maize lethal necrosis disease (MLND) emerged in Kenya (Wangai et al. 2012) and subsequently spread to Tanzania (FAO REOA 2013; Read, Featherston, et al. 2019a), Rwanda (Adams et al. 2014), the Democratic Republic of the Congo (Lukanda et al. 2014), Ethiopia (Mahuku, Wangai, et al. 2015), and Uganda (Kagoda et al. 2016). In 2017, crop losses in Africa of between 281 and 339 million USD were attributed to MLND (Pratt et al. 2017), suggesting that this new disease may be as economically devastating as MSD. MSV has recently been detected in plants expressing MLND symptoms, giving rise to the concern that MSV may play a role in MLND symptom severity (Redinbaugh and Stewart 2018). Based on climate suitability for the viruses that cause MLND and their insect vectors, it has been predicted that by 2050 MLND may spread to South Africa through Mozambique and/or Zimbabwe (Isabirye and Rwomushana 2016).

Considering that the most recent study of MSV genetic diversity and distribution in maize in South Africa was conducted on plants sampled ~20 years ago (Martin et al. 2001), the current status of MSV in the country is unknown. Thus, pre-empting the spread of MLND, the aim of this study was to determine the current genetic diversity and distribution of MSV in maize in South Africa as well as in the neighbouring maize-growing regions of Eswatini (formerly Swaziland) and the Maputo province of Mozambique.

3.2 Materials and methods

3.2.1 Re-evaluation of MSV type and subtype designation

All complete MSV genome sequences (> 2.6 kb) available on GenBank (National Center for Biotechnology Information, Bethesda, Maryland, USA) were downloaded, linearised to the same start position, and aligned using standard settings on CLC Genomics Workbench 20.0.2 (https://digitalinsights.qiagen.com). A CLC model test was performed, and a Neighbour-Joining tree assembled using the most suitable model available on the software,

the Kimura 80 model, and 1,000 bootstrap replicates. A CLC pairwise nucleotide identity matrix was assembled from the alignment and operational taxonomic units (OTUs) designated based on a subtype threshold of 98%, that was defined by Muhire et al. (2013). OTUs comprising sequences that represented previously recognised types and subtypes were annotated accordingly.

Where possible, two representatives of each designated OTU were aligned along with Digitaria streak virus (DSV) as an outgroup using MUSCLE on MEGA X (Kumar et al. 2018), and a Maximum Likelihood tree constructed using the model of best fit, Tamura-Nei model with Gamma distribution (5 rate categories), invariant sites, and 1,000 bootstrap replicates. The results obtained from the two methods were compared and OTUs representing distinct types and subtypes determined.

3.2.2 Sampling and nucleic acid extraction

Maize leaves with virus-like symptoms were collected along five maize grain transport routes identified by the project funders, the South African National Seed Organization (SANSOR) (Supplementary 3.1), including route 1 described in Chapter 2 (section 2.1). A total of 56 representatives of the samples expressing various MSV-like symptoms were selected for MSV genetic diversity analysis (Supplementary data 3.2). These samples originated from sampling sites located in four of South Africa's provinces, the Maputo province of Mozambique and across Eswatini, with the highest number of sites sampled from Eswatini, KwaZulu-Natal and Limpopo (Fig. 3.1). Nucleic acid extracts of suitable concentration and purity were prepared for each sample as described in Chapter 2 (section 2.1) (White et al. 2008). A polymerase chain reaction (PCR) that amplifies mRNA (180 bp) and DNA (270 bp) of the actin gene present in plants was used as a control to confirm the presence of DNA in the nucleic acid extracts of the selected plant samples as described in Chapter 2 (section 2.2) (van den Berg et al. 2004).

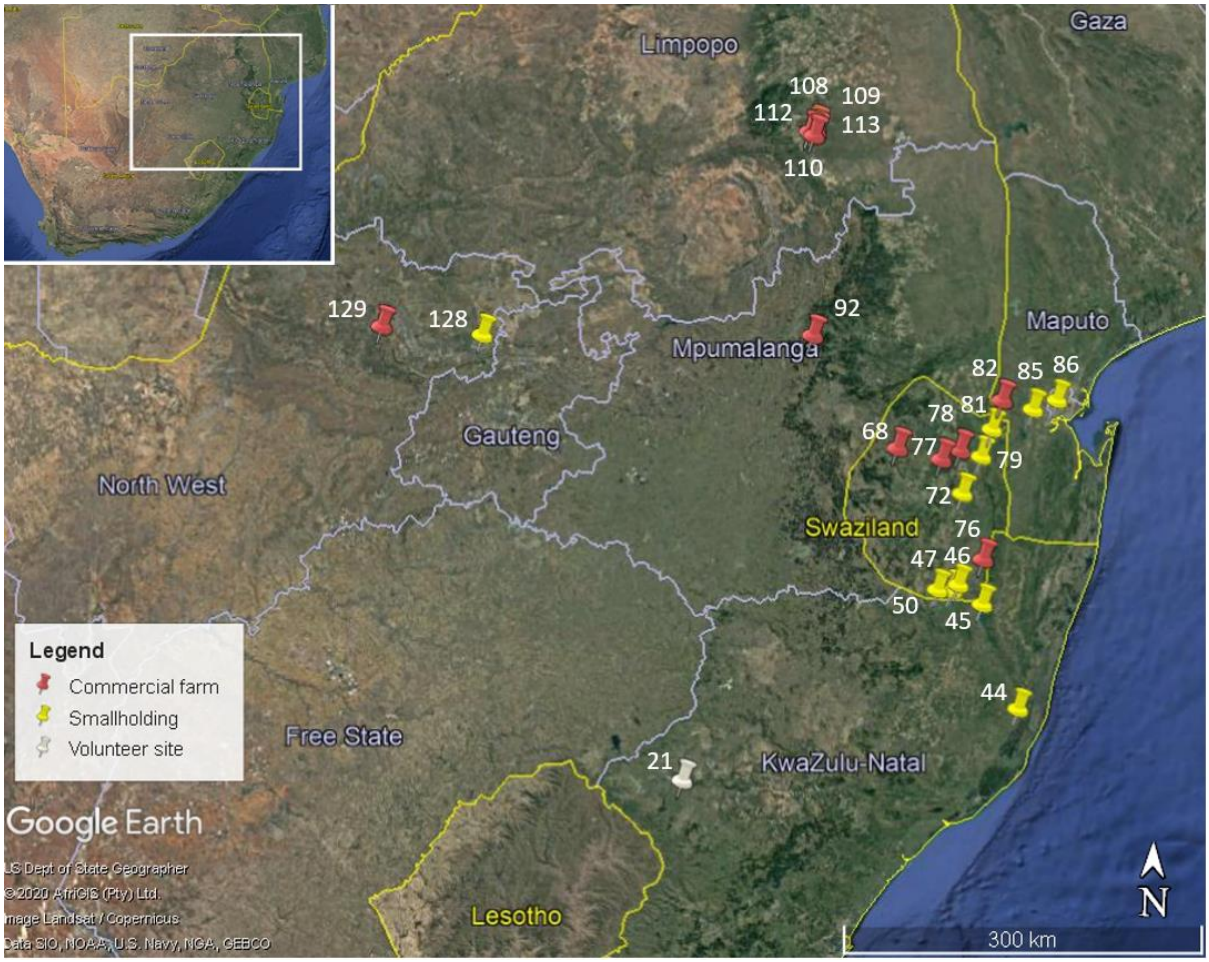


Fig. 3.1 Locations of sites where samples with maize streak virus (MSV) like symptoms were selected for genetic diversity analysis of the long intergenic region of the MSV genome. Adapted from Google Earth Pro 7.3.3.7786 (2015).

3.2.3 Analysis of MSV hypervariable region

3.2.3.1 PCR amplification

A set of primers (MSV F: 5′- CCA AAK DTC AGC TCC TCC G -3′; and MSV R: 5′- TTG GVC CGM VGA TGT ASA G -3′) with an expected amplicon size of ~1300 bp (Willment et al. 2001) was used to amplify a hypervariable region of the MSV genome spanning the conserved regulatory protein, replicase A (RepA) gene, hypervariable long intergenic region (LIR) and a portion of the movement protein (MP) gene. The hypervariable region of each of the selected samples was amplified using PCR. Bands of expected size were extracted, gel purified, and sent for bidirectional Sanger sequencing (Central Analytical Facility, Stellenbosch, ZA) as described in Chapter 2 (section 2.3).

The forward and reverse sequences were aligned using CLC Main Workbench 7.6.2 (https://digitalinsights.qiagen.com) and ambiguous bases resolved where possible. Base calls remained ambiguous when a second peak was at least 50% as high as the main peak. The consensus sequences were then extracted, and the identities confirmed as MSV using BLASTn similarity searches (National Center for Biotechnology Information, Bethesda, Maryland, USA).

3.2.3.2 Phylogenetic analysis

Sequence representatives of MSV types and subtypes determined from known sequences (section 2.1) were downloaded from GenBank, linearised, and the sequences rearranged to produce an unbroken sequence alignment spanning the hypervariable region. The sequences produced during this study for the MSV long intergenic regions and an outgroup (DSV) were aligned to the downloaded sequences using MUSCLE in MEGA X and the alignment trimmed to represent a cognate region. A Maximum Likelihood tree was constructed using the model of best fit, General Time Reversible model with Gamma distribution (5 rate categories), and 1,000 bootstrap replicates.

3.2.3.3 Phylogeographic analysis

Once type and subtypes of MSV variants present in the samples has been determined (section 2.3.2), relevant subtype representatives were then aligned using MUSCLE in MEGAX and the alignment imported into Network 10.0.1.1 (www.fluxus-engineering.com) (Bandelt et al. 1999) where a Median-Joining network was constructed. The nodes produced were colour-coded to distinguish between the MSV subtype references and the long intergenic region sequences produced during this study. The clustering obtained from the network was then compared to that of the Maximum Likelihood tree (section 2.3.2).

A Median-Joining network based on an alignment of only the long intergenic region sequences produced during this study was then constructed. The resulting nodes were colour-coded based on the province/region where each sample was collected. The subtype identities of the sequences were inferred from the Maximum Likelihood tree (constructed in section 2.3.2) and the respective clusters observed on the network labelled accordingly. The phylogeographic distribution of the MSV subtypes was then analysed.

3.2.4 Whole genome analysis

3.2.4.1 Next generation sequencing

To produce full genome sequence data for possibly divergent MSV variants present, samples with low nucleotide identity to known isolates, presence of nucleotide variations (ambiguous bases) within the sequences, and varying streak symptoms were selected for next generation sequencing (NGS). Libraries were prepared and RNA sequencing performed on an Illumina HiSeq 2500 instrument (2 x 125 bp paired-end reads) according to Shishkin et al. (2015). The resulting reads were then trimmed for quality as described previously (Chapter 2, section 2.5).

3.2.4.2 MSV type detection by reference mapping

Trimmed reads from each sample were multi-reference mapped to representatives of each MSV type (section 2.1) using an optimised length fraction of 1.00, similarity fraction of 0.94 and global mapping, with reads that mapped to multiple references ignored. The MSV type present in each sample was then determined based on the genome length mapped and depth of coverage obtained.

3.2.4.3 MSV subtype detection by reference mapping

The trimmed reads were then mapped to subtype (section 2.1) representatives of the relevant type detected using the same settings as before (section 2.4.2) but with an optimised similarity fraction of 0.98. The subtype present in each sample was then determined based on the genome length mapped and depth of coverage obtained.

3.2.4.4 MSV whole genome assembly

The trimmed reads from each sample were mapped to a reference of the single subtype that best suited the data using a length fraction of 0.9, similarity fraction of 0.9, and global mapping. Each sample's trimmed reads were subjected to *de novo* assembly with the following settings: minimum contig length: 500 nt; length fraction: 0.9; and similarity fraction: 0.9. Contigs obtained were then subjected to BLASTn (National Center for Biotechnology Information, Bethesda, Maryland, USA) for analysis. The MSV-like *de novo*-assembled contigs, reference-mapped sequences and (where possible) the long intergenic region sequences produced with PCR and Sanger sequencing (section 2.3.1) were aligned using standard settings and final consensus sequences of the MSV genomes extracted.

3.2.4.5 Phylogenetic analysis of MSV genomes

The complete MSV genome sequences were extracted and aligned to MSV type and subtype representatives along with DSV as an outgroup, using MUSCLE in MEGA X. The model of best fit Kimura 2-parameter model with Gamma distribution (5 rate categories), was then used to construct a Maximum Likelihood tree with 1,000 bootstrap replicates.

3.2.5 Other mastreviruses

Trimmed reads from samples with other viruses detected from the *de novo*-assembled contigs were reference mapped to suitable virus representatives using the same settings as described above (section 2.3.4). For each sample, the reference mapped sequence was aligned to the respective *de novo* contig(s) and the consensus sequences extracted, with "N" ambiguous bases used to fill sequence gaps where necessary. The sequences were then aligned to the genomes of other mastreviruses using MUSCLE in MEGA X and a phylogenetic tree constructed using the model of best fit, Kimura 2-parameter model with Gamma distribution (5 rate categories), and 1,000 bootstrap replicates.

3.3 Results

3.3.1 Reconstruction of known MSV types and subtypes

Using the pairwise identity threshold of 98%, 24 OTUs were identified, with 21 containing sequences previously characterised to at least type-level, as annotated on the Neighbour-Joining radial phylogeny (Fig. 3.2A). Three of the OTUs (3, 11 and 20) contained sequences that had only been identified to species-level previously. Representatives of two of the OTUs (19 and 21), contained sequences that had been identified as type MSV-C. The Maximum Likelihood tree supported the separation of these OTUs but suggested OTU 19 and 20 represented MSV-C while OTU 21 represented a previously uncharacterised type (Fig. 3.3) which was further supported by the identity matrix which showed that the sequences within this cluster shared less than 94% identity to all other sequences outside of this cluster.

Based on the clustering observed on the radial phylogeny as well as the previous identifications of some of the sequences within the clusters (Fig. 3.2A and B), 14 OTUs (1-9; 12-15; 20 and 21) appeared to represent possible subtypes of previously characterised MSV type A, B and C. OTUs 1-9 appeared to represent nine different MSV-A subtypes, three more than the six previously recognised (A_1-A_6), with sequences from three of

these (OTU 4-6) previously characterised as representatives of MSV-A_4 (Fig. 3.2B). In the Maximum Likelihood tree (Fig. 3.3) the two OTU 1 representatives clustered separately from one another, with GenBank accession HQ693332 clustering with MSV-A_1, while GenBank accession KJ699321 clustered separately from known MSV-A subtypes. Additionally, each of the OTU 4 (GenBank accession numbers: EU628573 and EU628576) and 5 (GenBank accession numbers: FJ882109 and KY618116) representatives clustered closely to one another but separate from other MSV-A subtypes (Fig. 3.3) suggesting that they represented a single previously uncharacterised subtype. Although MSV-A_5 appeared to represent its own subtype using the pairwise nucleotide identity approach, clustering produced by the Maximum Likelihood tree suggested MSV-A_5 closely enough with the MSV-A_1 representative to be considered a MSV-A_1 sublineage rather than a distinct subtype.

All of the other OTUs representing MSV-A subtypes identified based on pairwise identity (Fig. 3.2B) were supported by the Maximum Likelihood tree (Fig. 3.3), with OTU 3 (Accession number: KY618118) representing a previously unreported MSV-A subtype. In terms of the other MSV types, three OTUs, namely OTU 13-15 were identified as containing sequences that were previously characterised as representatives of MSV-B subtypes, namely B_3, B_1 and B_2, respectively (Fig. 3.2A), with OTU 11 clustering closely enough with MSV-B subtypes to fall under the MSV-B type but far enough to perhaps represent a novel subtype, and two OTUs (19 and 20) represented previously uncharacterised subtypes of MSV-C (Fig. 3.2A). These findings were also supported by the clustering observed in the Maximum Likelihood tree (Fig. 3.3).

Thus, for the purposes of this study, OTU 21 was defined as a novel type, tentatively named MSV-L; OTU 3 as a novel MSV-A subtype, tentatively named MSV-A_7; the OTU 1 representative that clustered separately from MSV-A_1 (GenBank accession: KJ699321) as a representative of a novel MSV-A subtype, tentatively named MSV-A_8; OTU 4 and 5 as jointly representing a novel MSV-A subtype, tentatively named MSV-A_9; OTU 11 as a novel subtype of MSV-B, tentatively named MSV-B_4; and OTU 19 and 20 representing two previously uncharacterised subtypes of MSV-C, tentatively named MSV-C_1 and MSV-C_2 (Fig. 3.3).

3.3.2 Analysis of hypervariable region

3.3.2.1 PCR amplification

The long intergenic region of MSV was successfully sequenced for each of the 56 selected samples with MSV-like symptoms. BLASTn similarity searches suggested that the amplicon sequences produced from the selected samples had between 96.80% and 98.85% nucleotide identity to one of only three reference sources: 42 to isolate MakD (MSV-A_5; GenBank accession: AJ012641), 12 to isolate VM (MSV-A_4; GenBank accession: AJ012637) and two to isolate ZA-Lad (subtype unclassified; GenBank accession: KY618103), with a range of between two and 23 ambiguous bases reported in 11 of the 56 sequences (Supplementary 3.3).

3.3.2.2 Phylogenetic analysis

The Maximum Likelihood phylogram suggests that the sequences produced for the long intergenic region of the selected MSV samples clustered most closely with MSV-A: 14 clustered most closely to MSV-A_4, while the other 42 clustered most closely to MSV-A_5 (Fig. 3.4). Similar clustering was observed from the networks constructed using Median-Joining (Fig. 3.5).

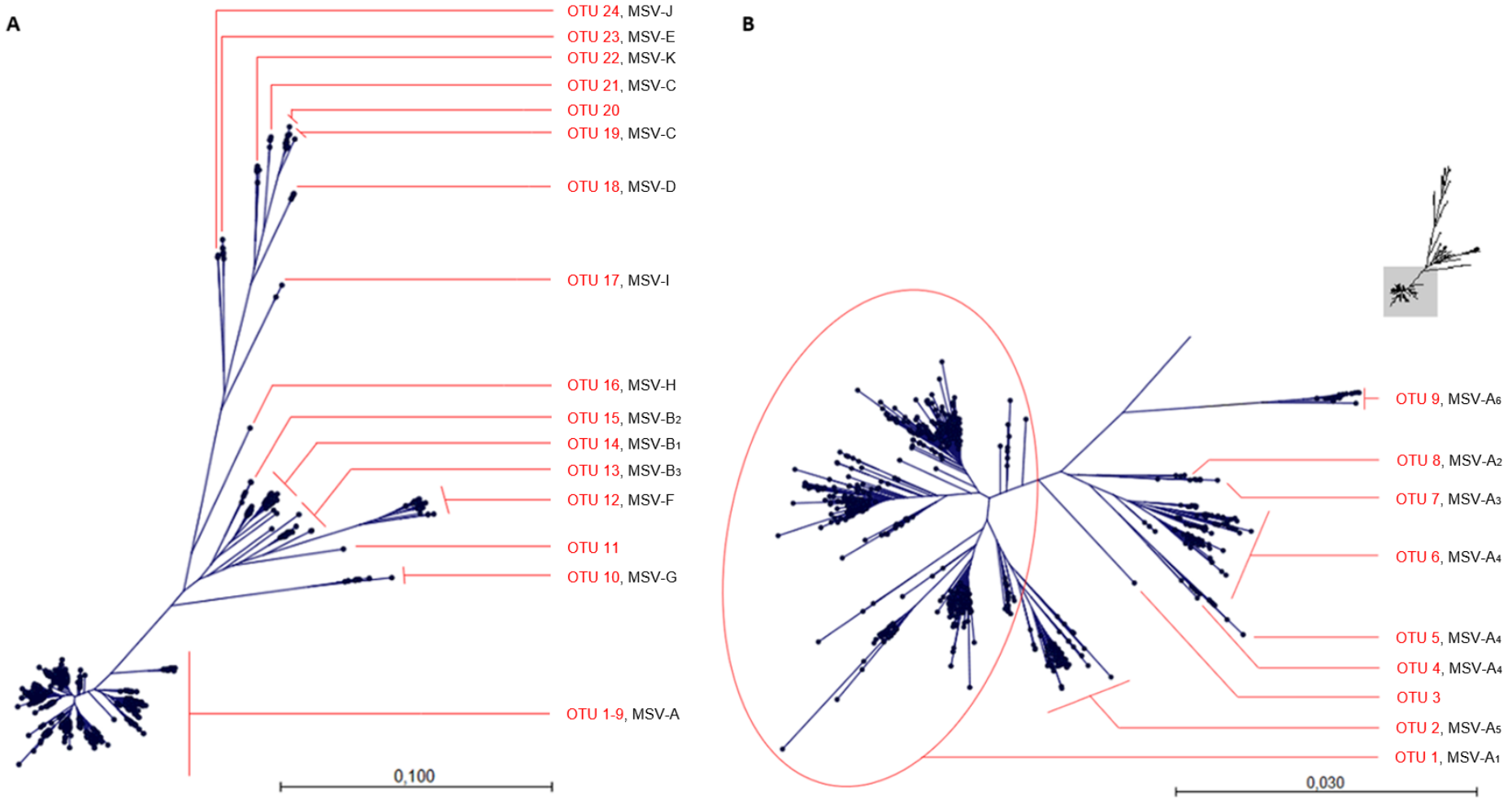


Fig. 3.2 Neighbour-Joining radial phylogeny based on complete genome sequences of all known maize streak virus (MSV) isolates (>2.6 kb) available on GenBank. The phylogram was constructed in CLC Genomics Workbench using the Kimura 80 model with 1,000 bootstrap replicates. There was a total of 911 sequences included in the final dataset. Designated operational taxonomic unit (OTU) annotations for different MSV subtypes sharing less than 98% pairwise nucleotide identity between clusters (excluding subtypes of MSV-A are depicted in **(A)** while MSV-A subtypes are depicted in **(B)**. The designated OTUs are displayed in red font, while clusters containing previously classified MSV types and subtypes are indicated in black font. Bars indicate the numbers of substitutions per site.

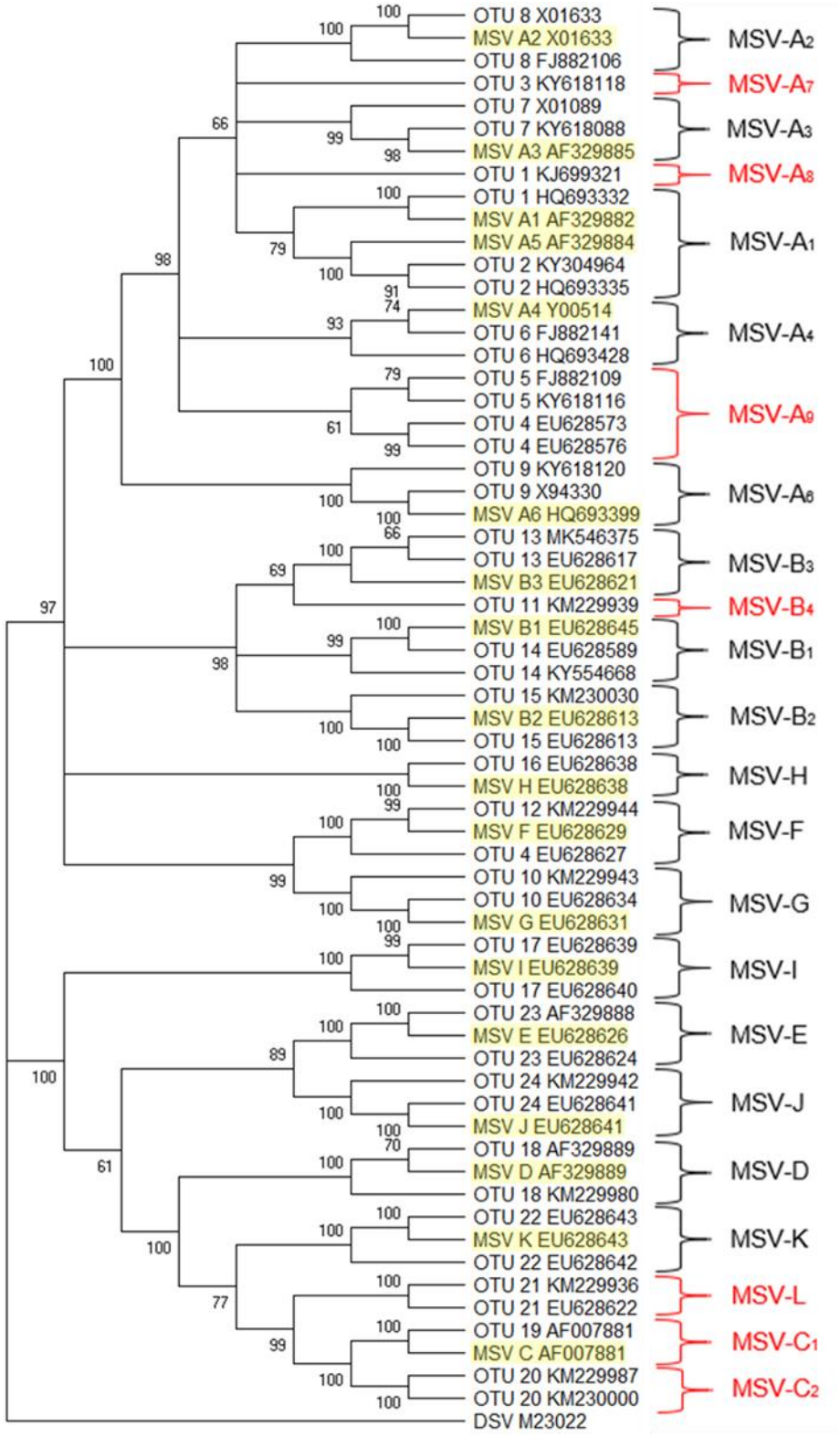


Fig. 3.3 Maximum Likelihood cladogram of operational taxonomic units (OTUs) described during this study against representatives of known maize streak virus (MSV) types and subtypes. The cladogram was constructed in MEGA X (Kumar et al. 2018) u sing the Tamura-Nei model with discrete Gamma distribution (5 rate categories), invariant sites, and 1,000 bootstrap replicates. Nodes with less than 60% bootstrap confidence were condensed. Previously classified MSV type and subtype representatives are highlighted in yellow and the outgroup, Digitaria streak virus (DSV). OTUs that represent potentially novel MSV types and subtypes are indicated in red, while those that clustered with previously characterised type and subtype references are annotated in black.

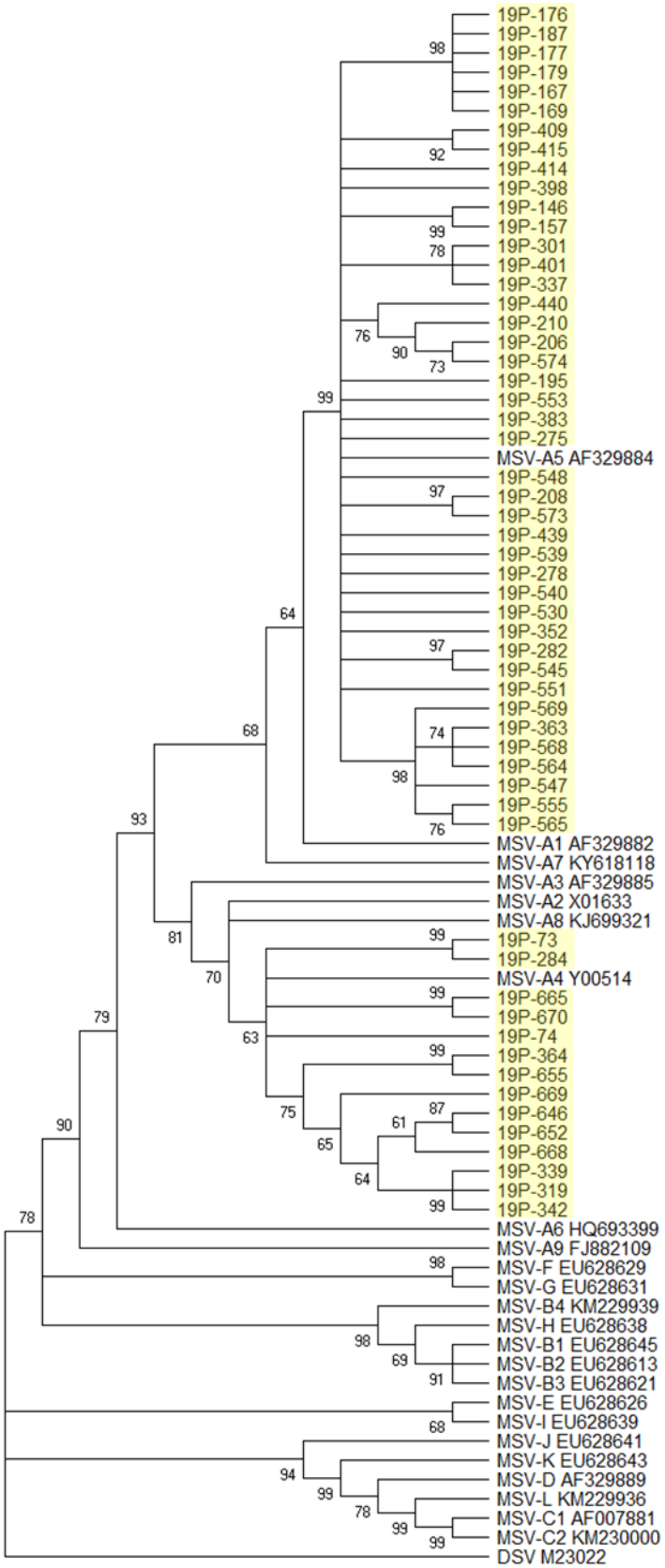


Fig. 3.4 Maximum Likelihood cladogram of the hypervariable long intergenic region of maize streak virus (MSV) sequenced from samples selected during this study against MSV type, subtype and sublineage representatives. The cladogram was constructed in MEGA X (Kumar et al. 2018) using the General Time Reversible model with discrete Gamma distribution (5 rate categories) and 1,000 bootstrap replicates. Nodes with less than 60% bootstrap confidence were condensed. Sequences produced during this study are highlighted in yellow.

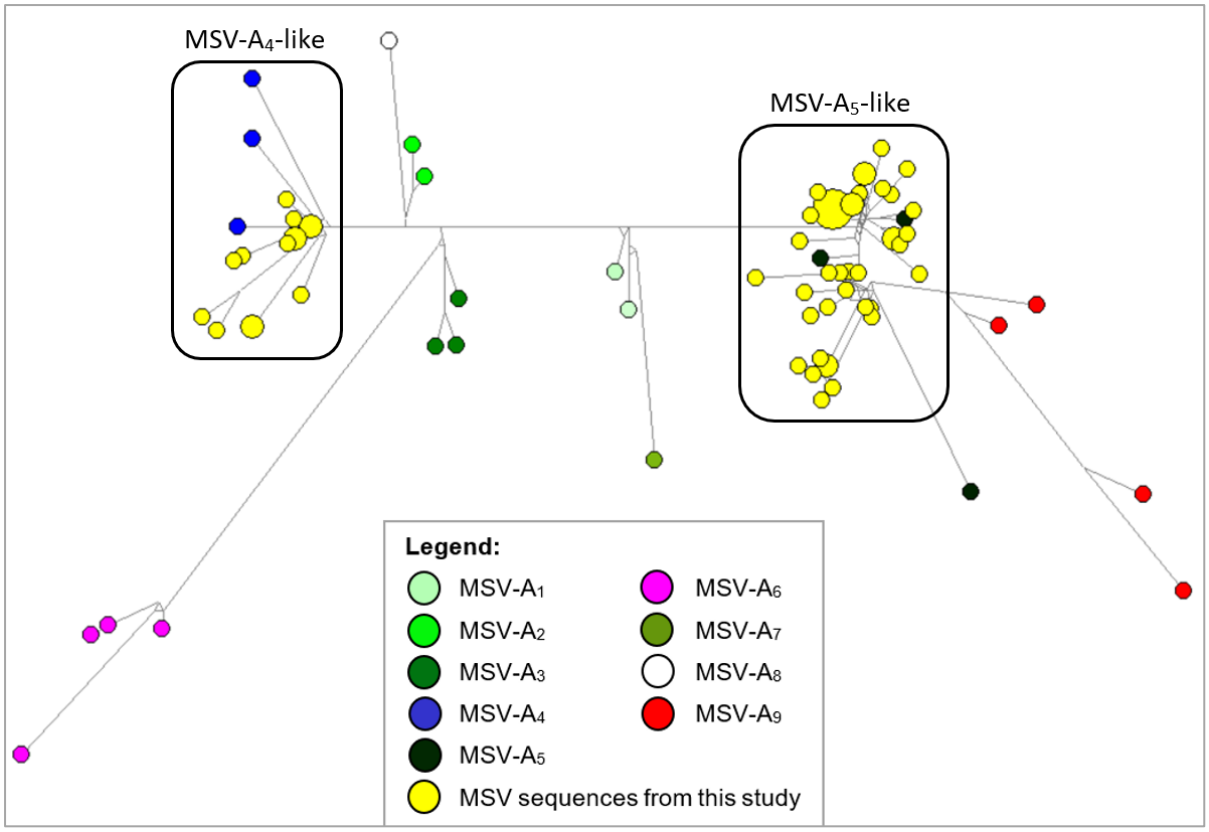


Fig. 3.5 Network of sequences produced during this study for the long intergentic region of maize streak virus (MSV), against representatives of recognised MSV-A subtypes (A_1-A_4 and A_6), MSV-A_5 (sublineage of MSV-A_1) and novel subtypes (A_{7-9}) defined during this study. Network created in Network 10 (Bandelt et al. 1999) using Median-Joining with node sizes proportional to the number of identical sequences represented.

3.3.2.3 Phylogeographic distribution

According to the phylogeographically colour-coded network of all sequences produced during this study (Fig. 3.6), only MSV-A_4-like variants were detected in the North West, only MSV-A_5-like variants were detected in Limpopo, Mpumalanga and Maputo (Mozambique), while both MSV-A_4- and MSV-A_5-like variants were detected in KwaZulu-Natal and Eswatini (see Supplementary 3.4 for nodes labelled with the specific sample accessions represented). These finds are however based on relatively low numbers of samples from the North West province and Mpumalanga.

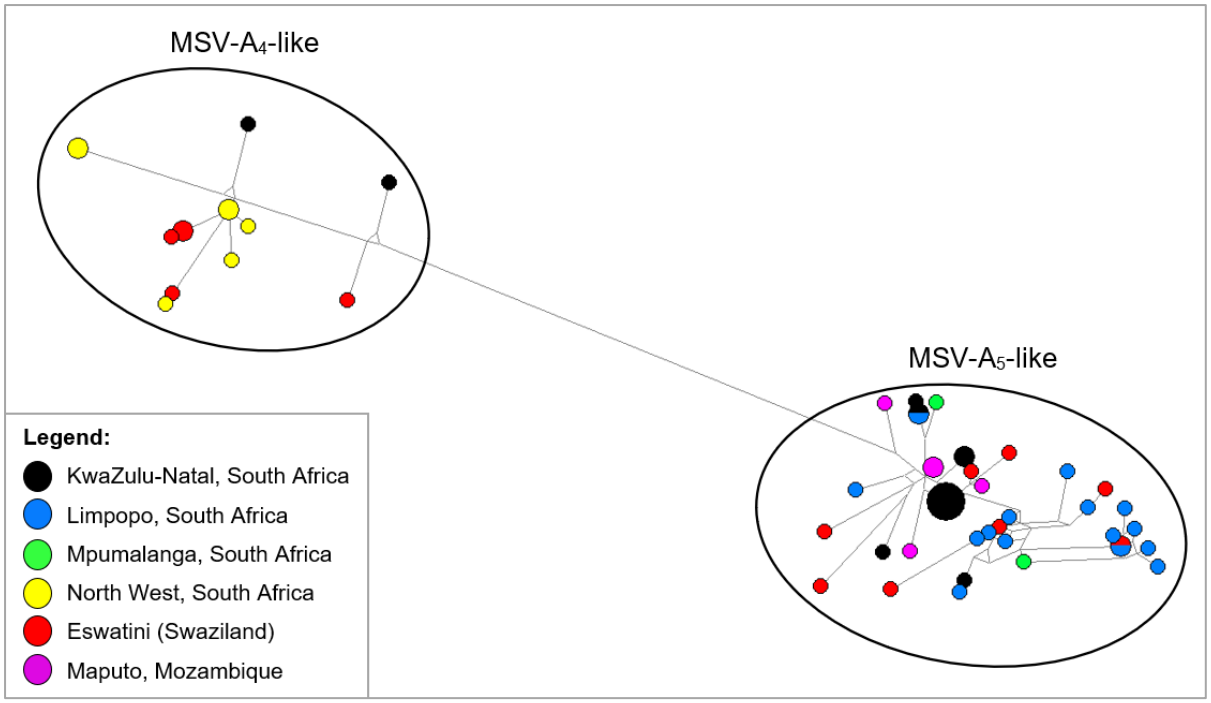


Fig. 3.6 Network of sequences produced during this study based on the long intergenic region of maize streak virus (MSV) to show geographic distribution of MSV-A_4-like and MSV-A_5-like variants. Network created in Network 10 (Bandelt et al. 1999) using Median-Joining with node sizes proportional to the number of identical sequences represented. Phylogenetic classifications based on clustering to reference sequences shown in Fig. 3.4.

3.3.3 Whole genome analysis

3.3.3.1 Next generation sequencing

Nucleic acid of four sample pools consisting of five plants each (19P-56 to -60, 19P-61 to -65, 19P-166 to -170 and 19P-171 to -175), and nine individual plants (19P-73, 19P-74, 19P-78, 19P-439, 19P-539, 19P-547, 19P-565, 19P-655 and 19P-670) were sent for NGS. Following quality trimming, a range of 601,913 to 4,069,383 reads remained per sample with an average length of 119 nt.

3.3.3.2 *De novo* assembly

Using the *de novo* assembly approach, 257-3,821 contigs were produced with a length range of 431-12,524 nt across the 13 different samples. Of these contigs, 15 had 98.05-99.90% identity to known MSV isolates and five had 97.40-99.13% identity to maize streak Reunion virus (MSRV) (Table 3.1). The MSV-like contigs ranged in length from 537 to 2,690 nt and average coverage depth of 15.32 to 1,366.12 X, while the MSRV-like contigs ranged in length from 548 to 1,574 nt and average coverage depth of 10.78-309.92 X.

Table 3.1 BLASTn hits of *de novo*-assembled contigs produced for 13 different samples to known isolates of maize streak virus (MSV) and maize streak Reunion virus (MSRV) (in bold font). *De novo* assembly performed on CLC Genomics Workbench 20.0.2 (https://digitalinsights.qiagen.com) with length and similarity fractions of 0.9. E-values for all BLASTn hits were 0.0 with query coverage of 98-100% except where indicated: * = 86% query coverage.

Sample/ sample pool	*De novo*-assembled contig		BLASTn hit			
	Contig length (nt)	Average coverage depth (X)	Identity (%)	Hit accession	Length (nt)	Accession description
19P-56 to -60	1741	36.62	99.89	HQ693446	2690	MSV-A_ZA_War6_Ta6-2008
19P-61 to -65	1162	15.32	99.66	HQ693445	2690	MSV-A_ZA_War5_Ta4-2008
19P-73	2690	95.40	99.07	FJ882095	2688	MSV-A_LS_Mal1_Les1_2005
19P-74	2519	27.20	99.72	KY618116	2688	MSV-A_ZA_Val_g524_2010
19P-78	1032	132.09	99.90	HQ693439	2690	MSV-A_ZA_War1_T5-2007
19P-166 to -170	916	55.35	99.56	KY304889	2687	MSV_KE_Ker5_K189B_10
19P-171 to -175	1654	24.83	98.70	MN428858	2687	MSV_Rwanda_20
	994	254.88	99.60	KY304889	2687	MSV_KE_Ker5_K189B_10
19P-439	**1278**	**201.28**	**98.27**	**MK329305**	**2879**	**MSRV-MV-252**
	824	98.87	99.62	KY618101	2688	MSV-A_ZA_Let_O85_1987
	578	**10.78**	**99.13**	**MK329305**	**2879**	**MSRV-MV-252**
19P-539	921	119.25	99.35	KY618100	2687	MSV-A_ZA_Let_O83_1987
	537	1366.12	99.63	HQ693353	2687	MSV-A_MZ_Map3_Moz24-2007
	1371	**64.04**	**98.06**	**MK329305**	**2879**	**MSRV-MV-252**
19P-547	2064	460.35	98.91	HQ693358	2687	MSV-A_MZ_Map8_Moz3-2007
	1476	**309.92**	**98.12***	**MK329305**	**2879**	**MSRV-MV-252**
	1574	**29.92**	**97.40**	**KT717933**	**2880**	**MSRV-YN**
19P-565	2658	256.13	98.05	MN428858	2687	MSV_Rwanda_20
19P-655	2647	113.44	99.44	EU152259	2690	MSV-SBlu
19P-670	1458	704.04	99.51	EU628575	2690	MSV-A4_Za-RosE-g131-2006

3.3.3.3 MSV type identification

Upon performing a multi-reference map against representatives of different MSV types, reads successfully covered between 82.5% and 99.9% of the MSV-A reference genome, while a range of just 4.4 to 23.9% genome coverage was obtained for other MSV type references (Table 3.2). This suggested that MSV-A was the dominant, if not only, type present in the selected samples.

Table 3.2 Percentages of genome lengths covered by trimmed next generation sequencing (NGS) reads mapped to different maize streak virus (MSV) type representatives for 13 samples. GenBank accession numbers of MSV type representatives were as follows: A: Y00514; B: EU628597; C: AF007881; D: AF329889; E: EU628626; F: EU628629; G: EU628631; H: EU628638; I: EU628639; J: EU628641; K: EU628643; L: EU628622. The multi-reference mapping was performed in CLC Genomics Workbench 20.0.2 (https://digitalinsights.qiagen.com) using a length fraction of 1.0, similarity fraction of 0.94, and global mapping.

Sample/sample pool	Percentage of genome length mapped per MSV type representatives											
	A	B	C	D	E	F	G	H	I	J	K	L
19P-56 to -60	99.9	8.6	-	-	-	-	-	-	-	-	-	-
19P-61 to -65	93.2	-	-	-	-	-	-	-	-	-	-	-
19P-73	99.7	7.0	-	-	-	-	9.7	8.2	-	-	-	-
19P-74	93.9	5.9	-	-	-	5.4	4.8	-	-	-	-	-
19P-78	93.1	4.4	-	-	-	-	-	-	-	-	-	-
19P-166 to -170	82.5	4.7	-	-	-	-	-	-	-	-	-	-
19P-171 to -175	93.2	12.1	-	-	-	-	-	9.3	-	-	-	-
19P-439	94.1	23.9	-	-	-	-	12.6	14.1	4.6	-	-	-
19P-539	94.6	17.9	-	-	-	-	11.0	10.5	-	-	-	-
19P-547	93.2	18.0	-	-	-	-	10.1	11.3	-	-	-	-
19P-565	94.0	8.8	-	-	-	-	12.5	4.5	4.5	-	-	-
19P-655	99.6	-	-	-	-	-	-	-	-	-	-	-
19P-670	99.8	17.9	-	-	-	7.8	-	-	-	-	-	-

3.3.3.4 MSV subtype identification

When reads were mapped to representatives of the eight MSV-A subtypes, the greatest genome length coverage was obtained for one of four subtypes (A_1, A_2, A_4 or A_9) per sample, with portions covered ranging from 19.2-56.5% across the 13 samples (Table 3.3A). However, the greatest average coverage depths were obtained for one of three subtypes (A_1, A_4 or A_6) per sample, with a range of 0.4-74.2 X (Table 3.3B). The reference mapping was then repeated using a references of broader subtype clusters.

Upon revision of the Maximum Likelihood tree that was used to designate the MSV-A subtypes (Fig. 3.3), the subtypes appeared to cluster into four larger groups. The one group included subtypes A_1 (including sublineage A_5), A_2, A_3, A_7, and A_8, while subtypes A_4, A_6 and A_9 individually comprised the other three groups. A single representative of each of the four groups was then used and the reference mapping repeated. From this mapping, seven samples produced the best mapping against the A_4 representative while six produced the best mapping against the A_5 representative (Table 3.4). The best genome length coverage ranged from 41.4-99.9% and average coverage depth of 5.7-250.1 X.

Table 3.3 Trimmed next generation sequencing (NGS) reads mapped to representatives of eight subtypes of maize streak virus (MSV) type A, described during this study, for 13 different samples/sample pools. GenBank accession numbers of MSV-A subtype references: A_1: AF329882; A_2: X01633; A_3: AF329885; A_4: Y00514; A_6: HQ693399; A_7: KY618118; A_8: KJ699321; and A_9: FJ882109. Results shown as **(A)** percentage of genome length covered, and **(B)** average coverage depth per reference sequence. The multi-reference mapping was performed in CLC Genomics Workbench 20.0.2 (https://digitalinsights.qiagen.com) using a length fraction of 1.0, similarity fraction 0.98 and global mapping. References mapped with the greatest length covered per sample highlighted in green.

A:

Sample/sample pool	Percentage of genome length covered per MSV-A subtype							
	A_1	A_2	A_3	A_4	A_6	A_7	A_8	A_9
19P-56 to -60	10.6	29.5	24.6	48.1	9.3	10.1	4.6	14.8
19P-61 to -65	0.0	17.4	16.9	38.8	4.7	5.3	0.0	13.2
19P-73	16.0	23.0	5.4	43.8	0.0	0.0	9.3	23.5
19P-74	0.0	11.4	3.9	27.9	5.3	8.4	4.6	34.0
19P-78	15.4	16.4	21.8	43.0	9.0	9.1	0.0	15.9
19P-166 to -170	19.2	4.4	4.4	0.0	0.0	4.4	0.0	8.9
19P-171 to -175	39.2	5.0	5.4	0.0	0.0	6.1	7.7	13.0
19P-439	50.4	9.0	19.7	19.1	12.3	15.6	8.7	27.5
19P-539	56.5	0.0	20.3	10.6	12.0	11.3	4.7	29.8
19P-547	53.3	0.0	15.6	4.2	11.7	15.3	8.7	31.9
19P-565	43.0	5.9	10.7	0.0	11.2	6.6	4.4	25.2
19P-655	4.1	14.2	13.4	48.6	5.0	14.9	17.0	17.4
19P-670	25.3	47.7	28.3	46.1	6.4	21.2	10.6	12.4

B:

Sample/sample pool	Average depth of coverage per MSV-A subtype (X)							
	A_1	A_2	A_3	A_4	A_6	A_7	A_8	A_9
19P-56 to -60	1.9	1.7	1.1	17.3	1.0	1.0	0.0	4.6
19P-61 to -65	0.0	0.6	1.1	3.7	0.1	0.1	0.0	0.6
19P-73	2.1	2.9	0.3	41.1	0.0	0.0	0.5	5.3
19P-74	0.0	0.5	0.0	9.0	0.3	0.2	0.1	3.3
19P-78	1.5	0.6	0.9	12.1	0.5	1.2	0.0	4.4
19P-166 to -170	0.4	0.1	0.2	0.0	0.0	0.0	0.0	0.2
19P-171 to -175	7.6	0.1	0.7	0.0	0.0	0.2	0.2	1.0
19P-439	21.7	0.1	1.7	2.7	15.8	4.6	0.5	4.7
19P-539	18.6	0.0	1.3	0.7	10.9	4.4	0.6	4.2
19P-547	13.0	0.0	1.7	0.0	13.6	2.4	0.6	5.3
19P-565	10.7	0.2	1.1	0.0	7.0	1.9	0.0	3.8
19P-655	0.3	2.2	2.1	22.5	1.2	1.6	0.8	9.6
19P-670	15.9	7.4	18.0	74.2	8.2	7.3	0.5	19.6

Table 3.4 Percentage of genome length covered, and average depth of coverage obtained for trimmed next generation sequencing (NGS) reads mapped to representatives of four main groups of maize streak virus (MSV) type A subtypes as discussed during this study for 13 different samples. GenBank accession numbers of MSV-A subtype representatives: A_4: Y00514; A_5: AF329884; A_6: HQ693399; and A_9: FJ882109. The multi-reference mapping was performed in CLC Genomics Workbench 20.0.2 (https://digitalinsights.qiagen.com) using a length fraction of 1.0, similarity fraction 0.98 and global mapping. References mapped with the greatest length covered and coverage depth per sample are highlighted in green.

Sample/sample pool	Percentage of genome length covered per MSV-A subtype				Average depth of coverage per MSV-A subtype (X)			
	A_4	A_5	A_6	A_9	A_4	A_5	A_6	A_9
19P-56 to -60	79.3	16.9	13.3	21.9	28.7	2.3	1.2	6.9
19P-61 to -65	65.5	4.4	9.0	21.8	5.7	0.0	0.3	1.0
19P-73	63.2	7.6	0.0	23.5	54.0	1.1	0.0	5.7
19P-74	41.4	10.2	9.7	34.0	11.9	0.3	0.3	3.3
19P-78	79.9	6.2	13.7	23.2	22.7	1.5	0.6	7.5
19P-166 to -170	4.3	67.8	0.0	9.7	0.0	7.7	0.0	0.5
19P-171 to -175	12.8	94.0	0.0	12.9	0.6	50.4	0.0	2.6
19P-439	34.2	99.1	12.8	22.4	4.7	250.1	15.9	10.8
19P-539	25.7	99.9	16.0	22.3	2.2	184.5	11.1	10.8
19P-547	24.7	92.1	12.1	27.7	2.7	139.4	13.9	10.1
19P-565	11.3	90.5	11.2	27.3	1.4	79.4	7.0	6.4
19P-655	72.2	17.5	9.8	29.1	34.1	2.3	2.0	14.5
19P-670	79.3	25.6	10.8	20.4	162.0	12.9	8.5	50.0

3.3.3.5 Whole genome assembly

Complete/near-complete genome sequences were produced from the individual reference mappings, spanning 89.9% to 100.1% of the reference genome lengths with average coverage depths of between 14.9 X and 647.4 X (Table 3.5). When reference-mapped and *de novo*-assembled sequences were aligned to the Sanger-sequenced long intergenic region sequences, consensus sequences spanning the entire MSV genome (2,687-2,690 nt) were produced for nine of the 13 samples, with near-complete genomes (2,552-2,648 nt) obtained for the remaining four samples (Table 3.6). Upon BLASTn analysis, the complete/near-complete genomes had 97.76 to 99.78% identity to known GenBank isolates (Table 3.6).

Table 3.5 Sequences produced from trimmed next generation sequencing (NGS) reads mapped to best suited of maize streak virus (MSV) type A cluster references for 13 different samples.

Sample/sample pool	Reference used: GenBank accession	Genome length covered (%)	Average coverage depth (X)
19P-56 to -60	MSV-A_4: Y00514	100.0	82.4
19P-61 to -65	MSV-A_4: Y00514	94.8	14.9
19P-73	MSV-A_4: Y00514	100.0	95.9
19P-74	MSV-A_4: Y00514	96.6	26.3
19P-78	MSV-A_4: Y00514	97.3	62.6
19P-166 to -170	MSV-A_5: AF329884	89.9	24.6
19P-171 to -175	MSV-A_5: AF329884	99.2	118.5
19P-439	MSV-A_5: AF329884	100.0	647.4
19P-539	MSV-A_5: AF329884	100.0	501.8
19P-547	MSV-A_5: AF329884	100.0	391.5
19P-565	MSV-A_5: AF329884	100.0	396.2
19P-655	MSV-A_4: Y00514	99.6	115.7
19P-670	MSV-A_4: Y00514	99.9	464.9

Table 3.6 BLASTn hits with best bit-scores of maize streak virus (MSV) genome consensus sequences produced from a combination of *de novo* assembly, reference mapping and Sanger-sequenced polymerase chain reaction (PCR) products of the long intergenic region (LIR) (where possible) of 13 different samples. Samples where no PCR-based LIR sequences were available are indicated by *. The query coverage for all BLASTn hits was 99%, and E-values 0.0.

Sample/sample pool	Genome length covered (nt)	BLASTn hit		
		GenBank accession	Identity (%)	Hit description
19P-56 to -60 *	2690	EU628574	99.78	MSV-A4 Za-RosB
19P-61 to -65 *	2552	EU628574	99.46	MSV-A4 Za-RosB
19P-73	2690	FJ882095	98.96	MSV-A4 LS-Mal1
19P-74	2690	FJ882134	97.76	MSV-A Za_BooB
19P-78 *	2618	EU628574	98.19	MSV-A4 Za-RosB
19P-166 to -170	2572	FJ882098	99.38	MSV-A MZ-Chi4
19P-171 to -175 *	2664	FJ882098	98.62	MSV-A MZ-Chi4
19P-439	2688	FJ882127	98.81	MSV-A ZA-Por1
19P-539	2687	FJ882127	98.84	MSV-A ZA-Por1
19P-547	2687	FJ882127	98.81	MSV-A ZA-Por1
19P-565	2687	FJ882127	98.69	MSV-A ZA-Por1
19P-655	2690	EU628574	99.33	MSV-A4 Za-RosB
19P-670	2687	EU628575	99.18	MSV-A4 Za-RosE

3.3.3.6 Genome-wide phylogenetic analysis

The Maximum Likelihood tree based on the complete MSV genome suggests that the dominant variants present in the selected samples were most closely related to representatives of MSV-A_4 (19P-56 to -60, 19P-73, 19P-74, 19P-670 and 19P-655) and the MSV-A_5 sublineage of MSV-A_1 (19P-439, 19P-539, 19P-547 and 19P-565) (Fig. 3.7).

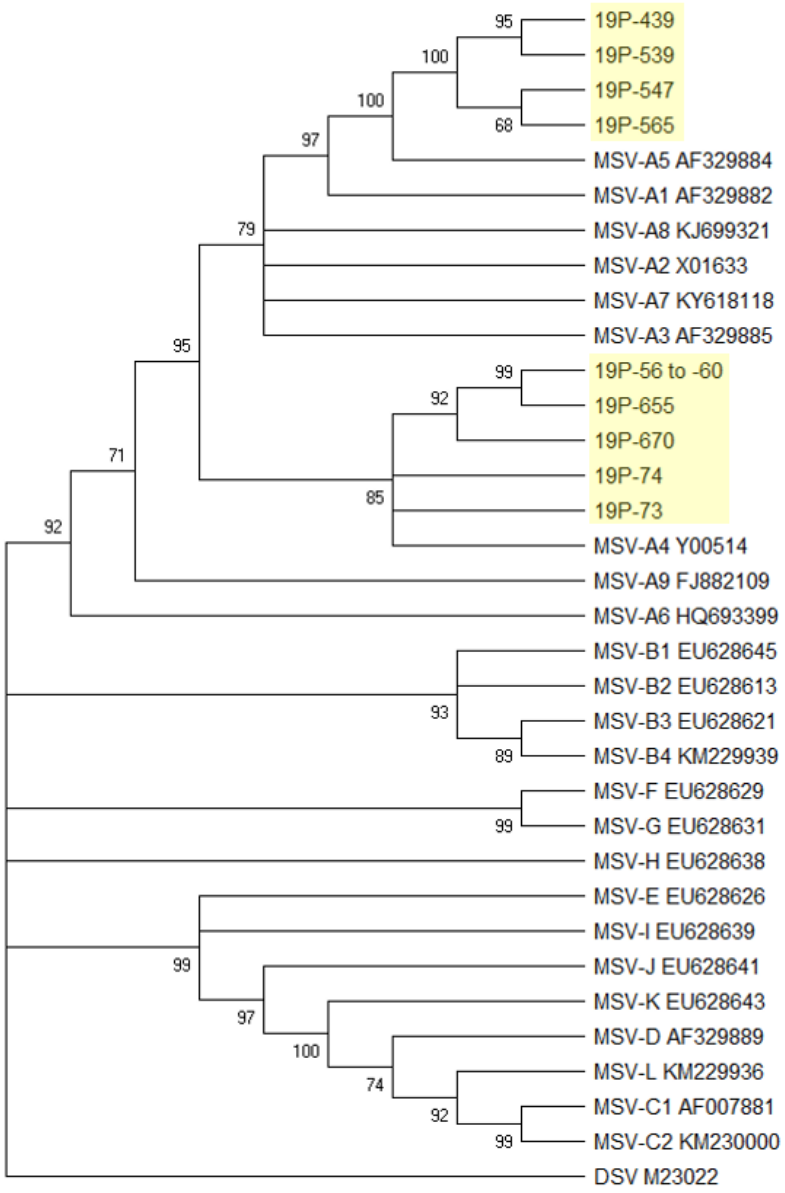


Fig. 3.7 Maximum Likelihood cladogram of the complete genome sequences of maize streak virus (MSV) type A variants produced during this study against MSV-A subtype reference sequences, along with MSV-A_5 (sublineage of MSV-A_1) and MSV-B as the outgroup. The cladogram was constructed in MEGA X (Kumar et al. 2018) with Kimura 2-parameter model with Gamma distribution (5 rate categories) and 1,000 bootstrap replicates. Nodes with less than 60% bootstrap confidence were condensed. Sequences produced during this study are highlighted in yellow.

3.3.4 Maize streak Reunion virus

Since the BLASTn analysis identified five of the *de novo*-assembled contigs from three of the samples (19P-439, 19P-539 and 19P-547) as having identity to maize streak Reunion virus (MSRV) (Table 3.1). Trimmed reads from these samples were then reference mapped to a MSRV representative (GenBank accession: KT717933), which produced one complete (19P-547: 2,879 nt) and two near-complete (19P-439: 2,863 nt; and 19P-539: 2,506 nt) genome sequences with an average coverage depth range of 37.6-138.6 X. BLASTn analysis of the consensus sequences produced by aligning both the *de novo*-assembled contigs and the reference-mapped sequences showed that the three sequences shared 99.31-99.90% pairwise nucleotide identity to one another and had 97.26-98.48%

identity to known MSRV isolate YN (Table 3.7). The MSRV consensus genome sequences clustered closely with previously identified MSRV variants available on GenBank (Fig. 3.8).

Table 3.7 BLASTn hits of maize streak Reunion virus (MSRV) genome consensus sequences produced from a combination of *de novo* assembly and reference mapping to a MSRV representative (GenBank accession: KT717933). The query coverage for all BLASTn hits was 99%, with E-values of 0.0.

Sample	Genome length covered (%)	BLASTn hit		
		Description	GenBank accession	Identity (%)
19P-439	99.4	MSRV isolate YN	KT717933	97.26
19P-539	87.0	MSRV isolate YN	KT717933	98.48
19P-547	100.0	MSRV isolate YN	KT717933	97.88

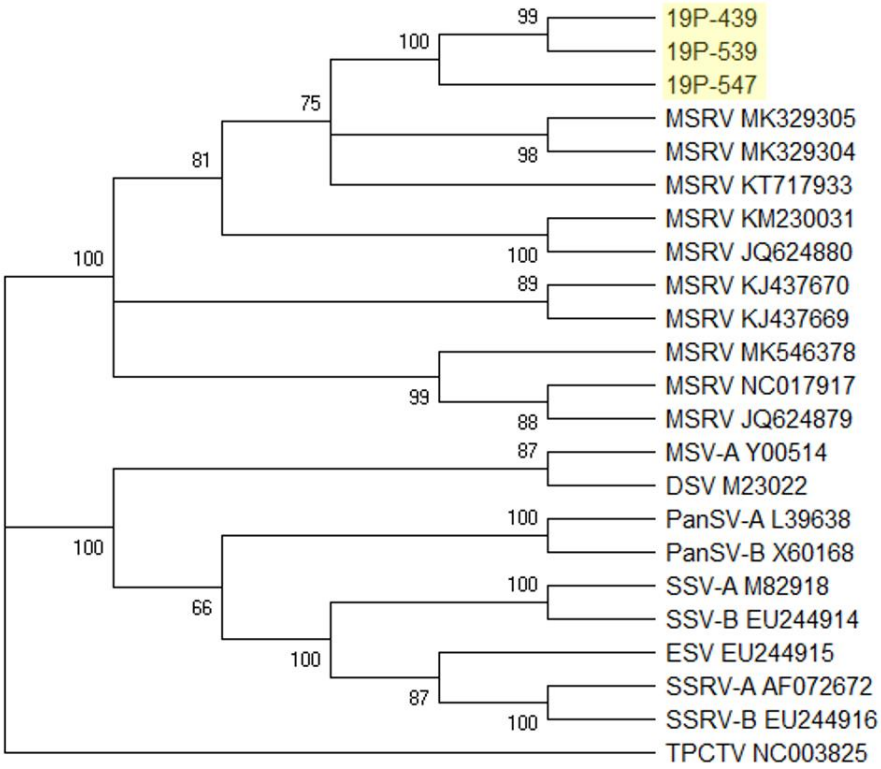


Fig. 3.8 Maximum Likelihood cladogram of near complete genome sequences of maize streak Reunion virus (MSRV) variants produced during this study against other members of the genus *Mastrevirus*. The cladogram was constructed in MEGA X (Kumar et al. 2018) with Kimura 2-parameter model with Gamma distribution (5 rate categories) and 1,000 bootstrap replicates. Nodes with less than 60% confidence were condensed. Tomato pseudo-curly top virus (TPCTV), a member of the genus *Topocuvirus*, family *Geminiviridae*, was used as an outgroup. Sequences produced during this study are highlighted in yellow.

3.4 Discussion

This study set out to determine the current genetic diversity of MSV in South Africa and the neighbouring maize-growing regions. A total of 56 samples with MSD-like symptoms from regions including KwaZulu-Natal, Mpumalanga, Limpopo and North West, South Africa, as well as Swaziland and Maputo, Mozambique were

selected from samples collected along five of South Africa's major maize grain transport routes. Before the genetic diversity of MSV present in the samples could be analysed, it was important to determine all possible types and subtypes that may exist currently, including and beyond what has been reported in literature to date.

Using a Neighbour-Joining tree of all known whole MSV genome sequences and a pairwise nucleotide identity matrix, subtypes were designated using a 98% threshold, as defined by Muhire et al. (2013), which specifies that genome sequences sharing at least 98% identity with one another be classified as the same subtype, and those sharing less than 98% identity to another sequence as representing a separate, distinct MSV subtype. A total of 24 OTUs were identified (Fig. 3.2A and B). To determine whether the OTUs represented distinct types/subtypes, a Maximum Likelihood tree was constructed using representatives of the OTUs and of previously defined types (MSV-A to -K) and subtypes (MSV-A_1 to -A_6 and MSV-B_1 to -B_3) (Fig. 3.3). When compared to the clustering produced by the Maximum Likelihood tree, discrepancies were observed.

The MSV-A_5 (OTU 2) clustered as a sublineage of MSV-A_1 (Fig. 3.3) as described by Owor et al. (2007) rather than a distinct subtype (Fig. 3.2B); one of the representatives of OTU 1 clustered separately of the other with clustered closely with the MSV-A_1 representative; and two other, separately defined OTUs (4 and 5) clustered closely enough to represent a single subtype when analysed with the Maximum Likelihood approach (Fig. 3.3). Overall, all previously identified types and subtypes were identified along with one new type, tentatively named MSV-L, three new MSV-A subtypes, tentatively named MSV-A_7, -A_8 and -A_9, one new MSV-B subtype, tentatively named MSV-B_4, and two new MSV-C subtypes, tentatively named MSV-C_1 and MSV-C_2 (Fig. 3.3).

The discrepancies observed between the percentage nucleotide identity approach and the clustering observed by the Maximum Likelihood tree may suggest that the subtype threshold is not highly robust and may present an over-simplified approach of subtype demarcation that may not reflect the true evolutionary relationships since this approach does not take evolutionary models, mutation rate differences across the genome, or the occurrence of synonymous/non-synonymous mutations into consideration. Thus, the clustering observed from the Maximum Likelihood tree, which takes evolutionary models, differing mutation rates and synonymous/non-synonymous mutations into account, was used to classify the OTUs in cases where such discrepancies existed.

The identified types and subtypes were then used as reference sequences to classify the MSV variants present in the maize samples collected in this study. The long intergenic region was used as a proxy to classify the variants to subtype level, to determine which samples may contain more diverged MSV variants, and for downstream phylogeographic analysis. Using a Maximum Likelihood tree with MSV type, subtype and sublineage (MSV-A_5) representatives, the long intergenic region sequences were shown to represent two subtypes of a single type, MSV-A_4 and MSV-A_1, with the MSV-A_1-like variants clustering closely with the MSV-A_5 sublineage representative (Fig. 3.4). The relationship between sequences based on sequence similarity produced by the Median-Joining network showed similar clustering (Fig. 3.5). When working with networks, it is important to take note of the limitations associated with this approach. Since network construction is based on an unrooted tree constructed using Median-Joining, networks are not suitable for accurate evolutionary or phylogenetic analyses (Kong et al. 2016). However, they are useful for observing relatedness based on the similarity observed between variants. Thus, all phylogenetic inferences made during this study were based on clustering patterns observed from Maximum Likelihood trees.

A total of 13 representatives believed to possibly contain diverged variants based on low BLASTn identities, network placement and different streak symptoms observed, were selected for further analysis. Although MSV is a DNA virus, NGS using RNA-Seq was used to analyse more of the genome as this was available at no monetary cost and the ability of RNA-Seq to detect DNA viruses had been discussed in literature previously (Pecman et al. 2017). Using a MSV type-focussed multi-reference map, the only dominant type present in the selected samples was determined as MSV-A (Table 3.2). Using a subtype-focussed approach and representatives of all eight subtypes discussed above (MSV-A_1 to -A_9, excluding A_5 which represented a sublineage of MSV-A_1), an inconclusive result was obtained, as the genome length coverage and average depth of coverage gave somewhat conflicting results (Table 3.3A and B). The genome length coverage observed indicated the dominant infections were that of A_1, A_2, A_4 or A_9, while the average depth of coverage indicated that A_1, A_4 or A_6 were the dominant infections.

Since MSV is known as a virus in which recombination occurs commonly, it is possible that the representatives used were too similar, thus the results may have been skewed as the "ignore" function that was used to create the mapping would have disregarded any reads that mapped to two or more of the references. To find representatives that would be less closely related and not recombinants of one another, the clustering of MSV-A subtypes into larger groups was analysed. The first cluster observed contained MSV-A_1, A_2, A_3, A_7, and A_8, while subtypes A_4, A_6 and A_9 formed three other independent clusters (Fig 3.3).

Thus, one representative of each cluster was used, with a MSV-A_5 representative used to represent the first cluster due to the variants in this study clustering closely with MSV-A_5 based on the long intergenic region sequences (Figs 3.4 and 3.5). This yielded far more conclusive results, with the highest genome length and average depth of coverage being obtained for the MSV-A_4 and MSV-A_5 references (Table 3.4), representing the dominant infections in the selected samples, and supporting the previous findings based on the long intergenic region sequences. Subtypes that had shorter regions covered by reads with lower coverage depths may have represented the presence of other MSV subtypes present at lower concentrations or be representative of quasispecies present within the plant.

For whole genome analysis, reads were mapped to the representative of the subtype cluster that best suited the sample with reduced stringencies to reduce the bias associated with reference mapping (Table 3.5). *De novo* assembly was also performed and the resulting sequences (Table 3.1), along with the reference mapped results (Table 3.5), and the long intergenic region sequences (Supplementary 3.3) were aligned to produce a consensus genome sequence (Table 3.6). Whole MSV genome sequences were produced for nine of the 13 samples (Table 3.6) and a Maximum Likelihood tree clustered the sequences with subtype MSV-A_4 and the MSV-A_5 sublineage of MSV-A_1 (Fig. 3.7), further supporting that these were the dominant MSV infections in the selected samples. Time constraints precluded confirmation of all sequences through Sanger sequencing of clones, but this will be done in future studies not only to confirm the sequences but also ensure that sequence chimeras are not produced as may happen if sequencing of overlapping primer regions were to be used. It should also be noted that since only samples expressing MSV-like symptoms were sampled during this study, and considering MSV-B to -K may not present symptoms on maize, it is possible that the sampling approach used in this study may have biased the MSV diversity study results.

Of the four samples where only partial MSV genomes were assembled, three of these represented pooled extracts from five individual plants each. Thus, it is possible that the pooling of multiple extracts may have reduced

the MSV concentration present, thus reducing the effective virus titre, and the number of relevant reads available for genome assembly, hence complete genome coverage was not possible. Another possibility is that the samples may have contained variants that were too diverged from the references used, however, due to the high identity of the partial genome sequences to other isolates available on GenBank (Table 3.6), this possibility seems unlikely. It should also be noted that there was no long intergenic region sequence data available for three of the four samples/sample pools (19P-61 to -65, 19P-78, and 19P-171 to -175). Thus, there was no long intergenic region sequence data to cover the gaps produced by NGS possibly due to the lack of reads available to cover that region as no mRNA would be produced from the intergenic regions.

The MSV phylogenetic trees constructed during this study were highly unresolved (Figs 3.1 and 3.8). This was not unexpected as the recombinant nature of viruses such as MSV make it impossible for the Maximum Likelihood models to determine accurate branching patterns between such recombinant variants (Gorbalenya and Lauber 2008). Such polytomy was observed for MSV-A_1, -A_2, -A_3, -A_7, and -A_8, and thus may lend further support for the clustering of these subtypes into a larger group as used for the reference mapping during this study (Table 3.4). For other viruses, phylogenetic trees have been constructed from amino acid sequences as they tend to be more conserved than nucleotide sequences (Read, Featherston, et al. 2019b, 2019c). However, that approach may not be suitable for MSV due to the common recombination events that occur, meaning that phylogenetic classification of MSV subtypes is highly complex. This may explain why different approaches and the use of different representative sequences may lead to different proposed clustering. It is proposed that after confirming the sequences produced during this study through Sanger sequencing of clones, an in-depth future study of recombination events within these sequences be pursued.

To determine the phylogeographic distribution of MSV in South Africa and the neighbouring maize-growing regions, a network of only the long intergenic region sequences produced during this study was assembled. By colour-coding the network based on the region(s) of origin of the sample(s) represented by each node, the phylogeographic distribution of the MSV variants could be determined. The two clusters were grouped into MSV-A_4-like and MSV-A_5-like (Fig. 3.6). Both MSV-A subtypes (MSV-A_4 and -A_5) were detected in the regions of KwaZulu-Natal and Eswatini (formerly Swaziland), while in the North West region only MSV-A_4 infections were detected, and in the Mpumalanga, Limpopo and Maputo, Mozambique region only MSV-A_5 infections were detected (Fig. 3.6). Of the two subtypes, MSV-A_5 was found to be the most common and widespread. Although MSV was reported in most provinces, no MSV-like symptoms or infections were reported in samples from the Free State (route 4, Supplementary 3.1). This was not surprising, however, as very few farms in that area had sown their maize crops by the time of the sampling due to the dry conditions South Africa faced in the early months of 2019 (Maluleke 2019). Maize growing regions in the Northern Cape were excluded from the survey for similar reasons.

Apart from MSV, the only other *Mastrevirus* species known to cause MSV-like symptoms on maize, MSRV (Pande et al. 2012), was also detected in the *de novo*-assembled contigs of three of the samples (19P-439, -539, and -547), sourced from a commercial farm in Emgwenya, Mpumalanga, and one commercial farm and one smallholding in Ofcolaco, Limpopo (Supplementary 3.2). These samples were subjected to NGS (Table 3.1). By combining the MSRV-like contigs with the results of reads mapped to a relevant MSRV reference, consensus sequences were obtained that represented the complete genome of one MSRV variant (sample 19P-547) and two near-complete genome sequences of two other variants (samples 19P-439 and -539) (Table 3.7). When analysed

against genome sequences of known mastreviruses, the three genomes were confirmed to cluster most closely with previously reported MSRV isolates. In future, overlapping primer sets will be designed over conserved regions of the MSRV genome to confirm the presence of the virus in South Africa and to confirm the genome sequences of the MSRV variants present. To date and according to genome sequences available on GenBank, MSRV has only been reported in plants from La Réunion, Nigeria, China, and, most recently, Ethiopia, and if confirmed, this would be the first report of this virus in South Africa.

In conclusion, despite MSV-resistant genotypes having been developed and commercially available in Africa for the last ~40 years (Soto et al. 1982), MSV continues to be widespread in South Africa, affecting both smallholder and commercial farms with MSV hotspots detected in Mkuze/Pongola (KwaZulu-Natal), and Ofcolaco (Limpopo) (Supplementary 3.2). A previously uncharacterised, novel MSV type was identified amongst the sequence data available on GenBank (tentatively named MSV-L), along with three previously uncharacterised, novel MSV-A subtypes (MSV-A_7, -A_8, and -A_9), one previously uncharacterised, novel MSV-B subtype (tentatively named MSV-B_4), and two previously uncharacterised, novel MSV-C subtypes (tentatively named MSV-C_1 and -C_2). Both the PCR-based approach of analysing the genetic diversity present within a hypervariable region and the NGS-based approach of whole genome analysis suggested that MSV type A was the dominant (if not only) type present in the selected samples with the variants detected clustering most closely with subtype A_4 and sublineage A_5 of subtype A_1. This study also highlighted the benefit of using a network-based approach simultaneously with phylogenetic trees to clearly visualise phylogeographic distributions of different virus variants, in this case different MSV-A subtypes. This study reports the current status of MSV infection in maize in South Africa and its neighbouring maize-growing regions as being similar to that reported in maize plants from 20 years ago (Martin et al. 2001). However, MSRV variants were detected in maize for the first time in South Africa during this study and will be formally reported once their presence and genome sequence(s) have been confirmed.

3.5 References

Adams, I. P., Harju, V. A., Hodges, T., Hany, U., Skelton, A., Rai, S., et al. (2014). First report of maize lethal necrosis disease in Rwanda. *New Disease Reports*, 29, 22.

Bandelt, H.-J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16(1), 37–48.

Bosque-Pérez, N. A. (2000). Eight decades of maize streak virus research. *Virus Research*, 71(1–2), 107–121.

FAO REOA. (2013). Maize lethal necrosis disease (MLND) - A snapshot. *GIEWS - Global Information and Early Warning System on Food and Agriculture*. http://www.fao.org/fileadmin/user_upload/emergencies/docs/MLND Snapshot_FINAL.pdf. Accessed 25 November 2020

Fuller, C. (1901). Mealic variegation. *First Report of the Government Entomologist Natal 1899-1900* (pp. 17–19). Pietermaritzburg: P. Davis & Sons, Government Printers.

Google Earth Pro 7.3.3.7786. (2015). South Africa, 26° 33’ 58.61’’S, 28° 51’ 37.52’’E, eye alt 829.08 km. Data SIO, NOAA, U.S. Navy, NGA, GEBCO. https://www.google.com/earth/index.html. Accessed 6 November 2020

Gorbalenya, A. E., & Lauber, C. (2008). Phylogeny of viruses. In B. W. J. Mahy & M. H. V. van Regenmortel (Eds.), *Encyclopedia of Virology* (3rd ed., pp. 125–129). Leiden: Elsevier Ltd.

Isabirye, B. E., & Rwomushana, I. (2016). Current and future potential distribution of maize chlorotic mottle virus and risk of maize lethal necrosis disease in Africa. *Journal of Crop Protection*, 5(2), 215–228.

Kagoda, F., Gidoi, R., & Isabirye, B. E. (2016). Status of maize lethal necrosis in eastern Uganda. *African Journal of Agricultural Research*, 11(8), 652–660.

Kong, S., Sánchez-Pacheco, S. J., & Murphy, R. W. (2016). On the use of median-joining networks in evolutionary biology. *Cladistics*, 32(6), 691–699.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549.

Lukanda, M., Owati, A., Ogunsanya, P., Valimunzigha, K., Katsongo, K., Ndemere, H., & Kumar, P. L. (2014). First report of maize chlorotic mottle virus infecting maize in the Democratic Republic of the Congo. *Plant Disease*, 98(10), 1448.

Mahuku, G., Wangai, A., Sadessa, K., Teklewold, A., Wegary, D., Ayalneh, D., et al. (2015). First report of maize chlorotic mottle virus and maize lethal necrosis on maize in Ethiopia. *Plant Disease*, 99(12), 1870–1870.

Maluleke, I. (2019). Looking back and ahead at the maize price. *Grain SA*. https://www.grainsa.co.za/looking-back-and-ahead-at-the-maize-price. Accessed 12 November 2020

Martin, D. P., Willment, J. A., Billharz, R., Velders, R., Odhiambo, B., Njuguna, J., et al. (2001). Sequence diversity and virulence in Zea mays of maize streak virus isolates. *Virology*. https://doi.org/10.1006/viro.2001.1075

Muhire, B., Martin, D. P., Brown, J. K., Navas-Castillo, J., Moriones, E., Zerbini, F. M., et al. (2013). A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). *Archives of Virology*, 158(6), 1411–1424.

Owor, B. E., Martin, D. P., Shepherd, D. N., Edema, R., Monjane, A. L., Rybicki, E. P., et al. (2007). Genetic analysis of maize streak virus isolates from Uganda reveals widespread distribution of a recombinant variant. *Journal of General Virology*. https://doi.org/10.1099/vir.0.83144-0

Pande, D., Kraberger, S., Lefeuvre, P., Lett, J.-M., Shepherd, D. N., Varsani, A., & Martin, D. P. (2012). A novel maize-infecting mastrevirus from La Réunion Island. *Archives of Virology*, 157(8), 1617–1621.

Pecman, A., Kutnjak, D., Gutiérrez-Aguirre, I., Adams, I., Fox, A., Boonham, N., & Ravnikar, M. (2017). Next generation sequencing for detection and discovery of plant viruses and viroids: Comparison of two approaches. *Frontiers in Microbiology*. https://www.frontiersin.org/article/10.3389/fmicb.2017.01998

Pratt, C. F., Constantine, K. L., & Murphy, S. T. (2017). Economic impacts of invasive alien species on African smallholder livelihoods. *Global Food Security*, 14(November 2016), 31–37.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019a). Characterization and detection of maize-associated pteridovirus (MaPV), infecting maize (Zea mays) in the Arusha region of Tanzania. *European Journal of Plant Pathology*, 154, 1165–1170.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019b). Diversity and distribution of maize-associated totivirus strains from Tanzania. *Virus Genes*, 55, 429–432.

Read, D. A., Featherston, J., Rees, D. J. G., Thompson, G. D., Roberts, R., Flett, B. C., et al. (2019c). Molecular characterization of Morogoro maize-associated virus, a nucleorhabdovirus detected in maize (Zea mays) in Tanzania. *Archives of Virology*, 164, 1711–1715.

Redinbaugh, M. G., & Stewart, L. R. (2018). Maize lethal necrosis: An emerging, synergistic viral disease. *Annual Review of Virology*, 5(August), 301–322.

Rybicki, E. P. (2015). A top ten list for economically important plant viruses. *Archives of Virology*, 160(1), 17–20.

Shishkin, A. A., Giannoukos, G., Kucukural, A., Ciulla, D., Busby, M., Surka, C., et al. (2015). Simultaneous generation of many RNA-seq libraries in a single reaction. *Nature Methods*, 12, 323–325.

Soto, P. E., Buddenhagen, I. W., & Asnani, V. L. (1982). Development of streak virus-resistant maize populations through improved challenge and selection methods. *Annals of Applied Biology*, 100(3), 539–546.

van den Berg, N., Crampton, B. G., Hein, I., Birch, P. R. J., & Berger, D. K. (2004). High-throughput screening of suppression subtractive hybridization cDNA libraries using DNA microarray analysis. *BioTechniques*, 37(5), 818–824.

Varsani, A., Shepherd, D. N., Monjane, A. L., Owor, B. E., Erdmann, J. B., Rybicki, E. P., et al. (2008). Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. *Journal of General Virology*. https://doi.org/10.1099/vir.0.2008/003590-0

Wangai, A. W., Redinbaugh, M. G., Kinyua, Z. M., Miano, D. W., Leley, P. K., Kasina, M., et al. (2012). First report of maize chlorotic mottle virus and maize lethal necrosis in Kenya. *Plant Disease*, 96(10), 1582.

White, E. J., Venter, M., Hiten, N. F., & Burger, J. T. (2008). Modified cetyltrimethylammonium bromide method improves robustness and versatility: The benchmark for plant RNA extraction. *Biotechnology Journal*, 3(11), 1424–1428.

Willment, J. A., Martin, D. P., & Rybicki, E. P. (2001). Analysis of the diversity of African streak mastreviruses using PCR-generated RFLPs and partial sequence data. *Journal of Veterinary Medical Science*, 93, 75–87.

Chapter 5: Conclusions

During this study, the presence and distribution of viruses implicated in the outbreak of maize lethal necrosis disease (MLND) in Tanzania, were determined along the major maize grain transport route of KwaZulu-Natal, South Africa. This was performed as a pre-emptive approach as MLND is predicted to spread to South Africa with KwaZulu-Natal being one of the provinces identified as high risk due to its climate being ideal for both MLND viruses and vectors to thrive. Samples with maize streak virus (MSV)-like symptoms from five of the maize-growing provinces in South Africa, as well as the neighbouring regions of Eswatini and Maputo, Mozambique, were also analysed for MSV genetic diversity and virus distribution since the current status of MSV in South Africa was unknown.

The presence of nine different maize viruses were detected using next generation sequencing (NGS), namely MSV, maize-associated pteridovirus (MaPV), Morogoro maize-associated virus (MMaV), two maize-associated totiviruses (MATV), maize stripe virus (MStV), two strains of Zea mays chrysovirus 1 (ZMCV1) and maize streak Reunion virus (MSRV), of which only MSV and MStV had been reported in maize in South Africa previously and with the report of MStV being unconfirmed. Assemblies of the NGS reads resulted in complete/near-complete genome sequences for all the viruses detected, except the two ZMCV1 strains that yielded only partial sequences. The presence of four of these viruses, namely MSV, MaPV, MMaV and MStV were also confirmed using polymerase chain reaction (PCR) and Sanger sequencing. Thus, the presence of MaPV, MMaV and MStV were reported for the first time in South Africa during this study, with the two MATV variants, MSRV, and the two ZMCV1 strains remaining as preliminary findings. The primary viruses associated with MLND, maize chlorotic mottle virus (MCMV) and maize-infecting potyvirids, were not detected in any of the samples analysed during this study.

The detection of some of the viruses with NGS but not PCR highlighted the benefits of NGS to detect novel viruses and emphasised the limitations of PCR-based virus detection. Sequence data available for MStV is currently extremely limited with only 23 sequences, including both partial and complete sequences of the various RNA segments, available on GenBank prior to this study. Complete/near-complete sequences for all five of MStV's RNA segments were produced during this study, with the RNA1 sequence potentially representing by far the most complete sequence produced for this segment to date. Prior to this study, the only partial sequence available for RNA1 spanned ~1.4 kb while the sequence produced during this study was ~9.0 kb in length. However, the sequences produced during this study only represent draft genome segment sequences, with further research using PCR and Sanger sequencing required for sequence confirmations.

Most of the viruses detected from the KwaZulu-Natal survey were sourced from sites closest to the Eswatini border. This raised concern as one of the predicted routes of MLND spread to South Africa is through Mozambique, which also shares a border with Eswatini. Most virus infections in KwaZulu-Natal were detected in maize from smallholder farms, where seed may be retained between seasons, providing a continual source of virus inoculum. MaPV and MMaV were detected in seemingly isolated cases, giving rise to the possibility that perhaps the possible vectors of these viruses may not be present in South Africa yet. However, further research into the incidence and distribution of these viruses across other maize-growing regions would be required to test

whether these are truly isolated cases. Since these viruses have only recently been discovered, future studies should also focus on possible symptoms associated with these viruses, their vectors and mode(s) of transmission, as well as the role they may play in MLND symptom expression, if any.

In terms of the genetic diversity and distribution of MSV in South Africa, the highest incidences of the viruses were found in KwaZulu-Natal, Limpopo and Eswatini, with similar genetic diversity to that reported 20 years ago, with the dominant virus type, MSV-A, and subtypes, MSV-A_1 (close relatives of the MSV-A_5 sublineage) and MSV-A_4 detected. However, the presence of MSRV detected using NGS in a MSV-infected sample, represents a novel finding as MSRV is the only other *Mastrevirus* species known to cause streak symptoms in maize, and had only been reported in a few areas outside of La Réunion previously, including Ethiopia, China and Nigeria.

Further research is currently underway to confirm the five RNA genome segment sequence of MStV and to develop confirmatory tests for MSRV. Other areas suggested for future research include determining possible symptoms associated with MaPV and MMaV, their vectors and mode(s) of transmission, as well as the role they and MSV may play in MLND symptom expression, if any; the presence and distribution of MaPV, MMaV, and MStV, as well as the other possible viruses preliminarily identified during this study, across other maize growing regions in South Africa; to determine whether co-infections of MSV and MStV additively affect the symptoms observed in local South African maize varieties; and to determine whether the MSV variants detected during this study may represent recombinants. Surveillance for MCMV and potyvirids in South African maize is encouraged to continue to detect regions where potyvirids may be present, and to enable the early detection of MCMV, should it be introduced into the country, to enable effective disease management procedures to be put in place.

Supplementary data

Supplementary 2.1 Site location, field type and symptoms observed for maize samples collected in KwaZulu-Natal with Global Positioning System (GPS) co-ordinates provided where available.

Date	Location	Site number	Longitude	Latitude	Altitude (m)	Field description	Plant accession	Observed symptoms
08/01/2019	Heidelberg	1	28°23'23.573"E	26°35'50.116"S	1621.335	Commercial field	19P-1	Somatic mutation
							19P-2	Darkish patches
							19P-3	Indistinct symptoms
							19P-4	Indistinct symptoms
							19P-5	Indistinct symptoms
		2	28°25'44.109"E	26°44'54.221"S	1551.99	Volunteer	19P-6	Little chlorotic speckles (fungus-like)
		3	28°27'06.420"E	26°47'31.295"S	1578.22	Commercial field	19P-7	Evenly distributed white dots
							19P-8	Reddening of margins and interveinal necrosis
	Villiers	4	28°38'31.454"E	27°03'55.086"S	1545.875	Commercial field	19P-9	Light green patch
							19P-10	Distinct necrotic spots
							19P-11	Distinct necrotic spots
							19P-12	Long chlorotic streaks
							19P-13	Long chlorotic streaks
							19P-14	Light green patch
	Villiers/Warden	5	28°44'06.678"E	27°13'57.995"S	1624.622	Commercial field	19P-15	Round chlorotic blotches (grass)
							19P-16	Little chlorotic speckles
		6	28°54'44.152"E	27°40'19.353"S	1638.368	Commercial field	19P-17	Long chlorotic streaks
							19P-18	Chlorotic patches
							19P-19	Chimera
							19P-20	Chimera and chlorotic streaks
							19P-21	Long chlorotic streaks
							19P-22	Mild mottle/speckled
	Warden	7	28°57'43.339"E	27°53'42.186"S	1699.728	Commercial field	19P-23	Red mid-rib
							19P-24	Vague chlorotic streaks

	Warden/Harrismith	8	29°00'29.511"E	28°02'20.078"S	1638.678	Commercial field	19P-25	Variable length broad streaks
							19P-26	Variable length broad streaks
							19P-27	Variable length broad streaks
		9	29°03'48.191"E	28°08'48.322"S	1696.29	Volunteer site	19P-28	Indistinct streaks
							19P-29	Chlorotic streaks along edge
							19P-30	Even chlorotic streaks
							19P-31	Red along leaf margins
							19P-32	Very even chlorotic interveinal areas
		10	29°05'28.376"E	28°11'30.409"S	1658.029	Commercial field	19P-33	Watermark blotches
							19P-34	Watermark blotches
							19P-35	Watermark blotches
							19P-36	Distinct edge chlorosis
							19P-37	Watermark blotches
							19P-38	Grey streaks (vertical patches)
							19P-39	Grey streaks (vertical patches)
							19P-40	Speckled
							19P-41	Speckled
	Harrismith/Swinburne	11	29°09'27.722"E	28°18'37.609"S	1670.797	Commercial field	19P-42	Brown necrotic streaks
							19P-43	Brown necrotic streaks
							19P-44	Vague interveinal yellowing
							19P-45	Fine white streaks
							19P-46	Fine white streaks
09/01/2019	Van Reenen	12	29°28'27.567"E	28°26'01.246"S	1155.829	Commercial field	No samples collected	No virus symptoms found
	Loskop	13	29°58'01.920"E	29°08'33.801"S	1524.719	Commercial field	19P-47	Mild chlorotic streak
							19P-48	Mild mottle
							19P-49	Mild mottle
	Mooiriver	14	29°49'17.415"E	29°01'17.187"S	1278.975	Commercial field	19P-50	Patch of flecks
							19P-51	Distinct chlorotic patch with green island
							19P-52	Even, feint chlorotic areas
	Estcourt	15	Not available	Not available	Not available	Commercial field	19P-53	A few chlorotic mottles
							19P-54	Chlorotic streaks along edge
							19P-55	Indistinct chlorotic patches

	16	Not available	Not available	Not available		19P-56	Chlorotic streaks
Estcourt/Winterton	17	29°34'04.870"E	28°56'01.039"S	1086.847	Commercial field	No samples collected	No virus symptoms found
Loskop	18	29°31'35.394"E	28°55'15.169"S	1158.482	Commercial field	19P-57	SCMV/MSV-like streaks
						19P-58	SCMV/MSV-like streaks
						19P-59	SCMV/MSV-like streaks (volunteer)
						19P-60	Multiple chimera
Winterton	19	Not available	Not available	Not available	Commercial field	19P-61	SCMV/MSV-like streaks
	20	Not available	Not available	Not available	Commercial field	19P-62	SCMV/MSV-like streaks
	21	29°32'43.0"E	28°53'34.0"S	Not available	Volunteers	19P-63	Interveinal chlorosis
						19P-64	Red vein banding
						19P-65	Red vein banding
						19P-66	Chlorotic marks
						19P-67	Chlorotic marks
						19P-68	Chlorotic marks
						19P-69	Indistinct mottle
						19P-70	Indistinct mottle
						19P-71	Indistinct mottle
						19P-72	SCMV/MSV-like streaks
						19P-73	SCMV/MSV-like streaks
						19P-74	SCMV/MSV-like streaks
						19P-75	SCMV/MSV-like streaks
	22	Not available	Not available	Not available	Volunteers	19P-76	Chlorosis on furling
						19P-77	Indistinct mottle
						19P-78	SCMV/MSV-like streaks
	23	29°32'15.175"E	28°49'43.065"S	1041.661	Commercial field	No samples collected	No virus symptoms found
	24	29°31'27.298"E	28°48'11.625"S	1035.509	Commercial field	19P-79	Systemic necrotic lesions (fungus-like)
						19P-80	Fine streaks laterally over leaf
	25	29°28'58.331"E	28°47'04.554"S	1093.213	Commercial field	No samples collected	No virus symptoms found
Bergville/Ladysmith	26	29°22'54.362"E	28°44'21.761"S	1139.542	Commercial field	No samples collected	No virus symptoms found
	27	29°21'09.372"E	28°42'21.726"S	1159.034	Commercial field	19P-81	Systemic necrotic lesions (fungus-like)
						19P-82	Systemic necrotic lesions (fungus-like)
						19P-83	Very distinct chlorotic dots and slashes

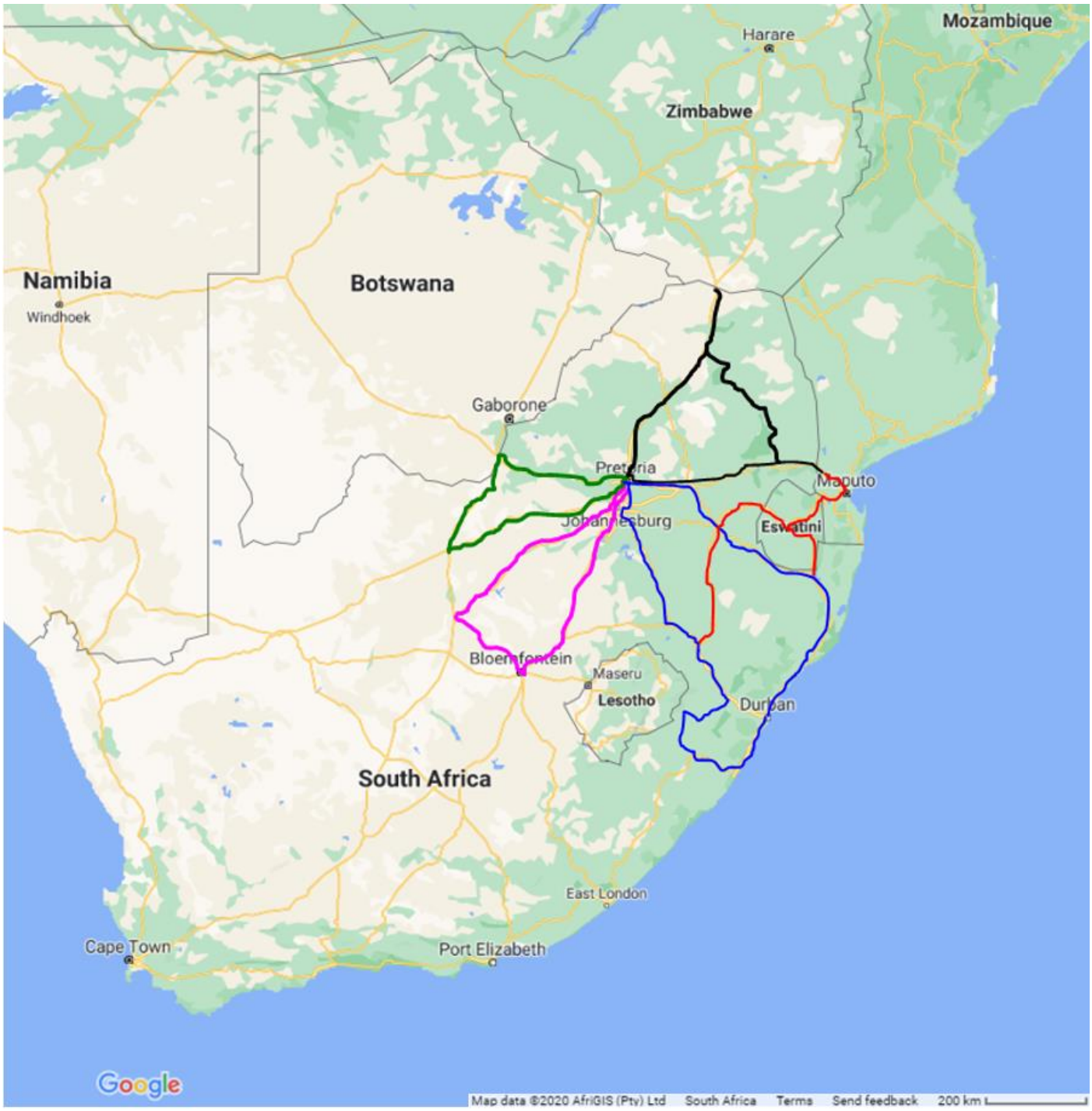
		28	29°26'11.998"E	28°36'50.442"S	1140.849	Commercial field	19P-84	Grey virus-like streaks
		29	29°31'56.299"E	28°35'29.791"S	1187.054	Commercial field	No samples collected	No virus symptoms found
	Mooiriver	30	29°58'20.324"E	29°09'19.917"S	1490.477	Commercial field	No samples collected	No virus symptoms found
	Mooiriver/Howick	31	30°10'00.513"E	29°25'10.669"S	1090.543	Commercial field	No samples collected	No virus symptoms found
10/01/2019	Umzimkulu	32	Not available	Not available	Not available	Commercial field	No samples collected	No Virus symptoms found
		33	29°55'07.697"E	30°21'26.006'S	988.904	Smallholding	19P-85	Indistinct chlorotic streaks
							19P-86	Indistinct chlorotic streaks
							19P-87	Interveinal chlorosis
							19P-88	Chlorotic streaks (fungus-like)
							19P-89	Even interveinal chlorosis (nutrient deficiency?)
							19P-90	Even interveinal chlorosis (nutrient deficiency?)
							19P-91	Even interveinal chlorosis (nutrient deficiency?)
							19P-92	Leaves with long chlorotic streaks
							19P-93	Chlorotic lesions on red background
							19P-94	Chlorotic streaks
							19P-95	Chlorotic lesions on red background
							19P-96	Chlorotic lesions on red background
							19P-97	Indistinct chlorotic speckles
		34	29°52'23.919"E	30°26'22.329"S	736.837	Smallholding	19P-98	Reddening of margins and interveinal necrosis
	Kokstad	35	29°32'34.445"E	30°31'13.623"S	1389.47	Commercial field	19P-99	Asymptomatic
							19P-100	Asymptomatic
		36	29°29'18.034"E	30°30'32.798"S	1307.597	Commercial field	19P-101	Small chlorotic flecks
							19P-102	Small chlorotic flecks in a patch
							19P-103	Large water-soaked chlorotic regions
							19P-104	Very clear interveinal chlorosis necrosis
							19P-105	Large water-soaked chlorotic regions
							19P-106	Long chlorotic streaks
							19P-107	Long chlorotic streaks
							19P-108	Long chlorotic streaks
							19P-109	Large chlorotic streaks
		37	29°22'16.678"E	30°32'54.523"S	1275.594	Commercial field	19P-110	Red mark along leaf margin
							19P-111	Mild mosaic

						19P-112	Very clear streaks along leaf margin
						19P-113	Chlorotic flecks in patches
						19P-114	Long chlorotic streaks
						19P-115	Very defined streaks
	38	29°28'21.829"E	30°38'36.575"S	1386.841	Smallholding	19P-116	Chlorotic streaks along leaf
						19P-117	Chlorotic patch across leaf
						19P-118	Various chlorotic streaks
						19P-119	Chlorotic speckles over leaf
						19P-120	Chlorotic streak along whole leaf
	39	29°29'30.279"E	30°33'10.967"S	1410.193	Commercial field	No samples collected	No virus symptoms found
Harding	40	29°48'21.080"E	30°33'07.000"S	984.323	Commercial field	19P-121	Mild mottle
						19P-122	Mild mottle
						19P-123	Mild mottle
						19P-124	Long chlorotic streaks (patch of chlorotic speckles)
						19P-125	Indistinct streaks
						19P-126	Chimera
Tongaat/Shaka's rock	41	31°11'26.388"E	29°32'09.627"S	69.79	Commercial field	19P-127	Indistinct chlorosis
						19P-128	Indistinct chlorosis
						19P-129	Indistinct chlorosis
Enyoni	42	31°32'34.578"E	29°09'00.790"S	66.903	Commercial field	No samples collected	No virus symptoms found
Mbosa	43	32°05'28.833"E	28°36'26.566"S	53.875	Smallholding	19P-130	Chlorotic streaks
						19P-131	Small chlorotic lesions
						19P-132	Indistinct chlorotic streaks
						19P-133	Small chlorotic streaks
						19P-134	Small chlorotic streaks
						19P-135	Indistinct chlorotic lesions
						19P-136	Small chlorotic lesions in patches
						19P-137	Indistinct chlorotic blotches
						19P-138	Small chlorotic lesions
						19P-139	Long chlorotic lesion
						19P-140	Chlorotic lesions
						19P-141	Chlorotic dot dash with streaks

							19P-142	General chlorosis of leaf
							19P-143	Small chlorotic lesions
							19P-144	Small chlorotic lesions
11/01/2019	Mfekayi	44	32°17'03.709"E	28°12'45.069"S	45.214	Smallholding	19P-145	General yellowing of leaf
							19P-146	Chlorotic streaks (MSV-like)
							19P-147	Chlorotic speckles (rust-like)
							19P-148	Virus-like chlorotic streaks; leaf yellowing
							19P-149	Virus-like chlorotic streaks; leaf yellowing
							19P-150	Single chlorotic streaks; leaf yellowing
							19P-151	Virus-like chlorotic streaks
							19P-152	Whole plant red leaves
							19P-153	Indistinct chlorotic streaks
							19P-154	Chlorotic lesions (rust-like)
							19P-155	Chlorotic lesions (rust-like)
							19P-156	Long chlorotic streaks
	Mkuze/Pongola	45	31°55'33.323"E	27°28'04.073"S	173.645	Smallholding	19P-157	SCMV/MSV-like streaks
							19P-158	Long chlorotic streaks
							19P-159	Chlorotic streaks
							19P-160	Long chlorotic streaks
							19P-161	General chlorosis with green islands
							19P-162	Chlorotic speckles
							19P-163	Small chlorotic lesions and streaks
							19P-164	General chlorosis with green islands
							19P-165	Chlorotic patch with green islands
							19P-166	Indistinct chlorotic streaks
							19P-167	SCMV/MSV-like streaks
		46	31°43'10.339"E	27°20'04.078"S	214.454	Smallholding	19P-168	SCMV/MSV-like streaks
							19P-169	SCMV/MSV-like streaks
							19P-170	SCMV/MSV-like streaks
							19P-171	Long chlorotic streaks
							19P-172	Watermarks (chlorotic)
							19P-173	SCMV/MSV-like streaks

							19P-174	SCMV/MSV-like streaks
							19P-175	SCMV/MSV-like streaks
							19P-176	SCMV/MSV-like streaks
							19P-177	SCMV/MSV-like streaks
							19P-178	Chlorotic lesions (rust-like)
							19P-179	SCMV/MSV-like streaks
							19P-180	General chlorosis with green islands
							19P-181	Chlorotic streaks
							19P-182	Long chlorotic streaks
							19P-183	Long chlorotic streaks
							19P-184	General chlorosis with green islands
							19P-185	Indistinct streaks
							19P-186	Water-soaked chlorotic lesions
							19P-187	SCMV/MSV-like streaks
	Pongola	47	31°40'07.517"E	27°22'44.284"S	229.307	Commercial field	19P-188	Indistinct chlorotic lesions (volunteer plant)
							19P-189	Virus-like long streaks
							19P-190	General chlorosis with green islands
							19P-191	Virus-like streaks
							19P-192	Chlorotic lesions
							19P-193	General chlorosis with green islands
							19P-194	lots of chlorotic spots
							19P-195	SCMV/MSV-like streaks
		48	31°34'40.830"E	27°23'04.546"S	346.358	Smallholding	19P-196	Water-soaked chlorotic lesions
							19P-197	Small chlorotic lesions in patch
							19P-198	Water-soaked chlorotic lesions
		49	31°34'13.619"E	27°23'03.820"S	366.732	Smallholding	19P-199	Chlorotic lesions along leaf margin
							19P-200	General chlorosis along leaf margin with green islands
							19P-201	Water-soaked chlorotic lesions
							19P-202	Chlorotic and necrotic lesions
							19P-203	Water-soaked chlorotic lesions
							19P-204	Chlorotic and necrotic lesions (rust-like)
		50	31°32'59.739"E	27°22'42.765"S	374.413	Smallholding	19P-205	Fine virus-like streaks

							19P-206	Fine virus-like streaks
							19P-207	Chlorotic lesions
							19P-208	Virus-like streaks
							19P-209	Small chlorotic speckles (fungus-like)
							19P-210	Interveinal chlorosis



Supplementary 3.1 The five major maize grain transport routes surveyed during this study. Figure legend: blue = route 1 (Chapter 2); red = route 2; black = route 3; fuchsia = route 4; and green = route 5. Image adapted from www.google.com/maps.

Supplementary 3.2 Locations of samples with MSV-like symptoms selected for genetic diversity analysis of the long intergenic region of the MSV genome. The names of countries where sampling occurred other than in South Africa are mentioned in brackets. Global positioning system (GPS) co-ordinates of the sampling sites have been provided where possible.

Plant accession	Sampling date	Site	Location	GPS co-ordinates	Site description
19P-73	09 01 2019	21	Winterton, KwaZulu-Natal	-	Volunteer site
19P-74	09 01 2019	21	Winterton, KwaZulu-Natal	-	Volunteer site
19P-146	11 01 2019	44	Mfekayi, KwaZulu-Natal	28°12'45.1"S 32°17'03.7"E	Smallholding

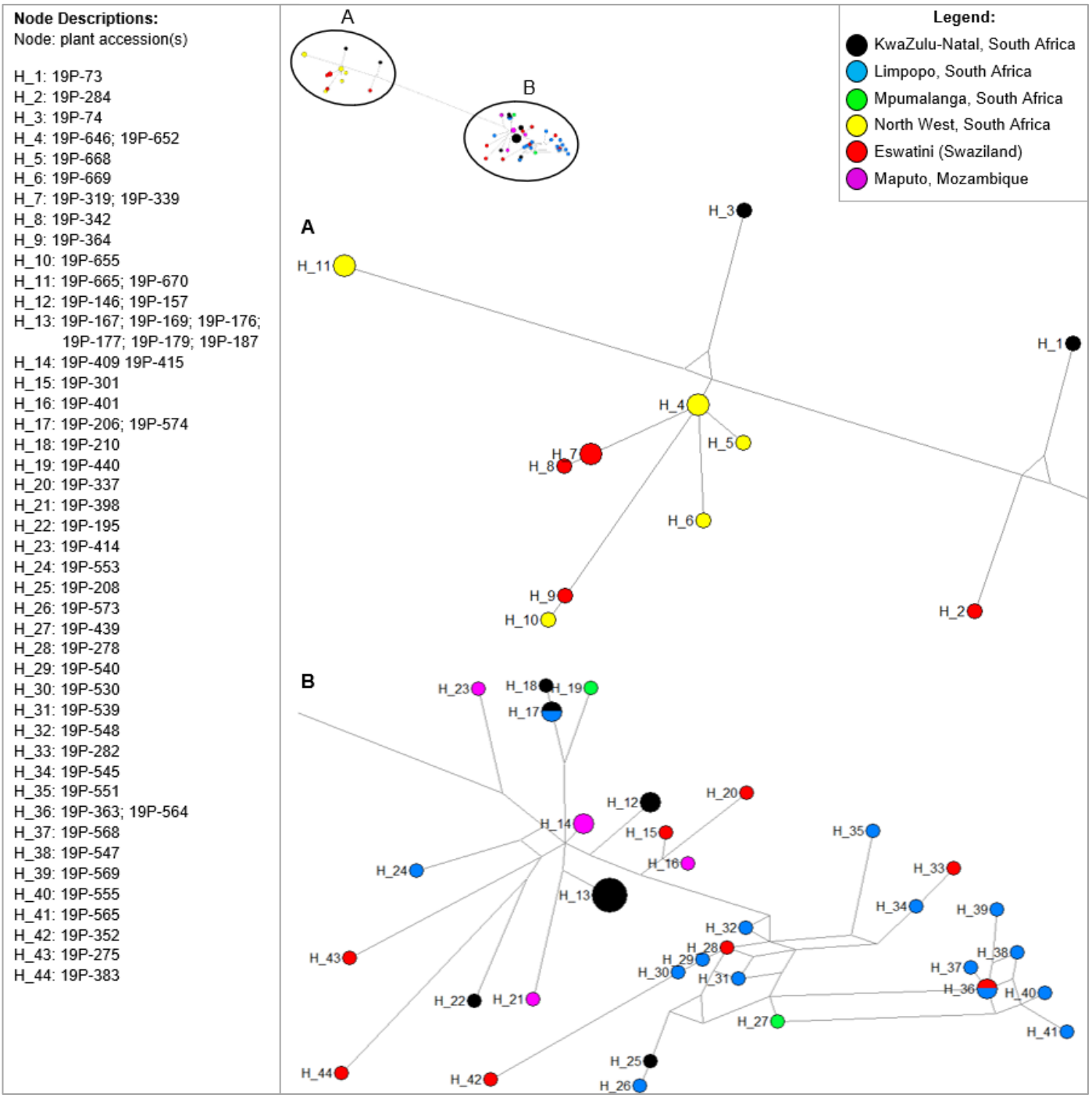
19P-157	11 01 2019	45	Mkuze, KwaZulu-Natal	27°28'04.1"S 31°55'33.3"E	Smallholding
19P-167	11 01 2019	45	Mkuze, KwaZulu-Natal	27°28'04.1"S 31°55'33.3"E	Smallholding
19P-169	11 01 2019	46	Mkuze, KwaZulu-Natal	27°20'04.1"S 31°43'10.3"E	Smallholding
19P-176	11 01 2019	46	Mkuze, KwaZulu-Natal	27°20'04.1"S 31°43'10.3"E	Smallholding
19P-177	11 01 2019	46	Mkuze, KwaZulu-Natal	27°20'04.1"S 31°43'10.3"E	Smallholding
19P-179	11 01 2019	46	Mkuze, KwaZulu-Natal	27°20'04.1"S 31°43'10.3"E	Smallholding
19P-187	11 01 2019	46	Mkuze, KwaZulu-Natal	27°20'04.1"S 31°43'10.3"E	Smallholding
19P-195	11 01 2019	47	Pongola, KwaZulu-Natal	27°22'44.3"S 31°40'07.5"E	Smallholding
19P-206	11 01 2019	50	Pongola, KwaZulu-Natal	27°22'42.8"S 31°32'59.7"E	Smallholding
19P-208	11 01 2019	50	Pongola, KwaZulu-Natal	27°22'42.8"S 31°32'59.7"E	Smallholding
19P-210	11 01 2019	50	Pongola, KwaZulu-Natal	27°22'42.8"S 31°32'59.7"E	Smallholding
19P-275	22 01 2019	68	Hhohho (Eswatini)	26°22'49.6"S 31°09'40.9"E	Commercial field
19P-278	22 01 2019	68	Hhohho (Eswatini)	26°22'49.6"S 31°09'40.9"E	Commercial field
19P-282	22 01 2019	68	Hhohho (Eswatini)	26°22'49.6"S 31°09'40.9"E	Commercial field
19P-284	22 01 2019	68	Hhohho (Eswatini)	26°22'49.6"S 31°09'40.9"E	Commercial field
19P-301	22 01 2019	72	Lubombo (Eswatini)	26°40'42.1"S 31°42'36.2"E	Smallholding
19P-319	22 01 2019	76	Lubombo (Eswatini)	27°08'37.9"S 31°54'45.9"E	Commercial field
19P-337	22 01 2019	77	Manzini (Eswatini)	26°26'34.1"S 31°31'27.8"E	Commercial field
19P-339	22 01 2019	78	Lubombo (Eswatini)	26°22'09.3"S 31°40'18.5"E	Commercial field
19P-342	22 01 2019	79	Lubombo (Eswatini)	26°24'17.1"S 31°50'49.9"E	Smallholding
19P-352	22 01 2019	79	Lubombo (Eswatini)	26°24'17.1"S 31°50'49.9"E	Smallholding
19P-363	22 01 2019	81	Simunye (Eswatini)	26°12'12.1"S 31°55'25.7"E	Smallholding
19P-364	22 01 2019	81	Simunye (Eswatini)	26°12'12.1"S 31°55'25.7"E	Smallholding
19P-383	23 01 2019	82	Lomahasha (Eswatini)	25°59'37.1"S 31°59'40.0"E	Commercial field
19P-398	23 01 2019	85	Congoana (Mozambique)	26°02'31.4"S 32°15'08.6"E	Smallholding
19P-401	23 01 2019	85	Congoana (Mozambique)	26°02'31.4"S 32°15'08.6"E	Smallholding
19P-409	23 01 2019	86	Matola (Mozambique)	25°57'50.2"S 32°26'48.5"E	Smallholding
19P-414	23 01 2019	86	Matola (Mozambique)	25°57'50.2"S 32°26'48.5"E	Smallholding
19P-415	23 01 2019	86	Matola (Mozambique)	25°57'50.2"S 32°26'48.5"E	Smallholding
19P-439	23 01 2019	92	Emgwenya, Mpumalanga	25°36'12.0"S 30°26'21.5"E	Commercial field
19P-440	23 01 2019	92	Emgwenya, Mpumalanga	25°36'12.0"S 30°26'21.5"E	Commercial field
19P-530	31 01 2019	108	Ofcolaco, Limpopo	24°05'07.9"S 30°24'02.5"E	Commercial field
19P-539	31 01 2019	108	Ofcolaco, Limpopo	24°05'07.9"S 30°24'02.5"E	Commercial field
19P-540	31 01 2019	108	Ofcolaco, Limpopo	24°05'07.9"S 30°24'02.5"E	Commercial field
19P-545	31 01 2019	109	Ofcolaco, Limpopo	24°06'43.7"S 30°23'16.2"E	Smallholding
19P-547	31 01 2019	109	Ofcolaco, Limpopo	24°06'43.7"S 30°23'16.2"E	Smallholding
19P-548	31 01 2019	109	Ofcolaco, Limpopo	24°06'43.7"S 30°23'16.2"E	Smallholding
19P-551	31 01 2019	110	Ofcolaco, Limpopo	24°09'26.0"S 30°23'27.5"E	Commercial field
19P-553	31 01 2019	110	Ofcolaco, Limpopo	24°09'26.0"S 30°23'27.5"E	Commercial field
19P-555	31 01 2019	110	Ofcolaco, Limpopo	24°09'26.0"S 30°23'27.5"E	Commercial field
19P-564	31 01 2019	112	Ofcolaco, Limpopo	24°07'37.7"S 30°24'08.1"E	Commercial field
19P-565	31 01 2019	112	Ofcolaco, Limpopo	24°07'37.7"S 30°24'08.1"E	Commercial field
19P-568	31 01 2019	112	Ofcolaco, Limpopo	24°07'37.7"S 30°24'08.1"E	Commercial field
19P-569	31 01 2019	113	Ofcolaco, Limpopo	24°07'51.2"S 30°21'21.2"E	Commercial field
19P-573	31 01 2019	113	Ofcolaco, Limpopo	24°07'51.2"S 30°21'21.2"E	Commercial field
19P-574	31 01 2019	113	Ofcolaco, Limpopo	24°07'51.2"S 30°21'21.2"E	Commercial field
19P-646	27 02 2019	128	Brits, North West	25°40'48.7"S 27°48'59.4"E	Smallholding
19P-652	27 02 2019	128	Brits, North West	25°40'48.7"S 27°48'59.4"E	Smallholding

19P-655	27 02 2019	128	Brits, North West	25°40'48.7"S 27°48'59.4"E	Smallholding
19P-665	27 02 2019	129	Koster, North West	25°37'57.1"S 27°00'47.3"E	Commercial field
19P-668	27 02 2019	129	Koster, North West	25°37'57.1"S 27°00'47.3"E	Commercial field
19P-669	27 02 2019	129	Koster, North West	25°37'57.1"S 27°00'47.3"E	Commercial field
19P-670	27 02 2019	129	Koster, North West	25°37'57.1"S 27°00'47.3"E	Commercial field

Supplementary 3.3 BLASTn analysis of bidirectional Sanger sequencing results from polymerase chain reaction products of a hypervariable region of the maize streak virus (MSV) genome. All hits had an E-value of 0.0.

Plant accession	BLASTn hit description	Query coverage (%)	Nucleotide identity (%)	GenBank accession	Ambiguous bases
19P-73	MSV isolate VM	100	97.95	AJ012637	-
19P-74	MSV isolate VM	100	96.80	AJ012637	23
19P-146	MSV isolate MakD	100	98.76	AJ012641	-
19P-157	MSV isolate MakD	100	98.76	AJ012641	-
19P-167	MSV isolate MakD	100	98.68	AJ012641	-
19P-169	MSV isolate MakD	100	98.68	AJ012641	-
19P-176	MSV isolate MakD	100	98.60	AJ012641	-
19P-177	MSV isolate MakD	100	98.60	AJ012641	-
19P-179	MSV isolate MakD	100	98.68	AJ012641	-
19P-187	MSV isolate MakD	100	98.68	AJ012641	-
19P-195	MSV isolate MakD	100	98.11	AJ012641	10
19P-206	MSV isolate MakD	100	98.60	AJ012641	2
19P-208	MSV isolate MakD	100	98.11	AJ012641	11
19P-210	MSV isolate MakD	100	98.19	AJ012641	14
19P-275	MSV isolate MakD	100	98.68	AJ012641	-
19P-278	MSV isolate MakD	100	98.85	AJ012641	
19P-282	MSV isolate MakD	100	98.27	AJ012641	-
19P-284	MSV isolate VM	100	98.11	AJ012637	
19P-301	MSV isolate MakD	100	98.85	AJ012641	-
19P-319	MSV isolate VM	100	97.53	AJ012637	6
19P-337	MSV isolate MakD	100	98.60	AJ012641	-
19P-339	MSV isolate VM	100	97.86	AJ012637	-
19P-342	MSV isolate VM	100	97.70	AJ012637	-
19P-352	MSV isolate MakD	100	97.94	AJ012641	-
19P-363	MSV isolate MakD	100	98.19	AJ012641	-
19P-364	MSV isolate VM	99	97.37	AJ012637	4
19P-383	MSV isolate MakD	99	98.51	AJ012641	-
19P-398	MSV isolate MakD	100	98.68	AJ012641	-
19P-401	MSV isolate MakD	100	98.77	AJ012641	-
19P-409	MSV isolate MakD	100	98.85	AJ012641	-
19P-414	MSV isolate MakD	100	98.85	AJ012641	-
19P-415	MSV isolate MakD	99	98.68	AJ012641	-
19P-439	MSV isolate MakD	100	98.11	AJ012641	10
19P-440	MSV isolate MakD	100	98.60	AJ012641	-
19P-530	MSV isolate MakD	100	98.52	AJ012641	-
19P-539	MSV isolate MakD	100	97.86	AJ012641	13
19P-540	MSV isolate MakD	100	98.76	AJ012641	-
19P-545	MSV isolate MakD	100	98.43	AJ012641	-
19P-547	MSV isolate MakD	100	98.02	AJ012641	6
19P-548	MSV isolate MakD	100	98.35	AJ012641	11
19P-551	MSV isolate MakD	100	98.52	AJ012641	-
19P-553	MSV isolate MakD	100	98.60	AJ012641	-
19P-555	MSV isolate MakD	100	97.94	AJ012641	-
19P-564	MSV isolate MakD	100	98.19	AJ012641	-
19P-565	MSV isolate MakD	100	97.86	AJ012641	-
19P-568	MSV isolate MakD	100	97.94	AJ012641	-
19P-569	MSV isolate MakD	100	97.94	AJ012641	-
19P-573	MSV isolate MakD	100	98.60	AJ012641	-
19P-574	MSV isolate MakD	100	98.68	AJ012641	-
19P-646	MSV isolate VM	100	98.19	AJ012637	-

19P-652	MSV isolate VM	100	98.11	AJ012637	-
19P-655	MSV isolate VM	100	97.45	AJ012637	-
19P-665	MSV isolate ZA_Lad_g521_2010	100	98.53	KY618103	-
19P-668	MSV isolate VM	100	98.03	AJ012637	-
19P-669	MSV isolate VM	100	97.86	AJ012637	-
19P-670	MSV isolate ZA_Lad_g521_2010	100	98.53	KY618103	-



Supplementary 3.4 Phylogeographic distribution of maize streak virus (MSV) of the long intergenic region sequences produced during this study to show geographic distribution of (A) MSV-A_4-like variants, and (B) MSV-A_5-like variants. Network created in Network 10 (Bandelt et al. 1999) using Median-Joining and standard settings with node size proportional to the number of identical sequences represented.